



UNIVERSIDADE FEDERAL DO PARÁ
NÚCLEO DE PESQUISA EM ONCOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ONCOLOGIA E CIÊNCIAS MÉDICAS

**EXPRESSÃO DIFERENCIAL DE GENES REGULADOS PELO
MYC EM LINHAGENS DE CÂNCER GÁSTRICO**

Carla Mariana Ferreira Pessoa

BELÉM-PA

2018



UNIVERSIDADE FEDERAL DO PARÁ
NÚCLEO DE PESQUISA EM ONCOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ONCOLOGIA E CIÊNCIAS MÉDICAS

**EXPRESSÃO DIFERENCIAL DE GENES REGULADOS PELO
MYC EM LINHAGENS DE CÂNCER GÁSTRICO**

Autor: Carla Mariana Ferreira Pessoa

Orientador: Prof. Dr. Rommel Mario Rodríguez Burbano

Dissertação apresentada ao Programa de Pós-Graduação em Oncologia e Ciências Médicas, área de concentração: Oncologia, do Núcleo de Pesquisas em Oncologia da Universidade Federal do Pará, como requisito para a obtenção do título de Mestre em Oncologia e Ciências Médicas.

BELÉM-PA

2018



UNIVERSIDADE FEDERAL DO PARÁ
NÚCLEO DE PESQUISA EM ONCOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ONCOLOGIA E CIÊNCIAS MÉDICAS

**EXPRESSÃO DIFERENCIAL DE GENES REGULADOS PELO
MYC EM LINHAGENS DE CÂNCER GÁSTRICO**

Banca Examinadora:

Prof. Dr. Rommel Mario Rodríguez Burbano
Professor da Universidade Federal do Pará (UFPA)
Orientador

Prof. Dr. André Salim Khayat
Professor da Universidade Federal do Pará (UFPA)
Membro

Dra. Helem Ferreira Ribeiro
Universidade da Amazônia (Unama)
Membro

Prof. Dr. Paulo Pimentel de Assumpção
Professor da Universidade Federal do Pará (UFPA)
Membro

BELÉM-PA

2018

DEDICATÓRIA

*Ao professor e amigo Dr. Carlos Rocha, por
todas as contribuições acadêmicas e pessoais;
Aos meus pais Carmem e Francisco por todo o
amor e dedicação incondicionais.*

AGRADECIMENTOS

Agradeço a Deus, primeiramente por sua bondade infinita e proteção constante.

Aos meus pais por serem os melhores parceiros dessa vida, por todo amor e dedicação dispensados, especialmente à minha mãe, meu amor maior.

Às minhas irmãs e minhas sobrinhas, minhas companheiras de todas as horas.

Ao Jamerson Carvalho, meu companheiro e motivador nessa caminhada.

Ao Prof. Dr. Rommel Burbano por sua orientação e parceria no desenvolvimento deste documento.

Ao amigo Professor Carlos Rocha, minha eterna gratidão por estar sempre ao meu lado, por todas as inúmeras contribuições desde o estudo para a seleção até a finalização deste trabalho.

Ao Jersey, minha gratidão por toda a sua gentileza e atenção dedicados a mim.

Às parceiras de estudo Núbia Rabelo e Carla Sant'Anna.

Aos amigos Jorge, Nina, Alessandra e Laís muito obrigada pela parceria ao longo desses anos.

À família LCH que me acolheu com tanto carinho e respeito, serei eternamente grata por todo o aprendizado e por todos os momentos que compartilhamos juntos. Especialmente aos amigos Jorge, Emerson, Laís, Tatiane, Lorena, Emanuel, Jackeline, aos professores Marcelo, Carol e Adriana, e a todos que tive a satisfação de conhecer.

À minha orientadora de bancada, Ingrid Ramos, gratidão eterna por todas as valiosas colaborações e pelo apoio constante em todos os momentos.

À CAPES pelo apoio financeiro.

RESUMO

MYC é um oncogene responsável por crescimento celular excessivo no câncer, permitindo a ativação transcricional de genes envolvidos na regulação do ciclo celular, metabolismo e apoptose, sendo geralmente superexpressado no Câncer Gástrico (CG). Utilizando siRNA e Sequenciamento de Nova Geração (NGS), identificamos Expressão Diferencial de Genes (DEGs) regulados por *MYC* em três linhagens celulares brasileiras de CG representados pelos subtipos histológicos difuso, intestinal e metastático, e posteriormente integramos esses dados com um enriquecimento gênico computacional com a ferramenta GSEA (*Gene Set Enrichment Analysis*). Identificamos um total de 5.471 DEGs com uma correlação alta (80%). O silenciamento do *MYC* por siRNA, nas linhagens celulares de CG dos tipos difuso e metastático causou o aumento da quantidade de DEGs com expressão diminuída, enquanto na linhagem do tipo intestinal exibiu uma maior quantidade de DEGs com um perfil de expressão aumentada. A partir do enriquecimento gênico, utilizando nossas amostras sequenciadas comparando com a coleção *hallmark gene sets*, foram encontrados 11 conjuntos significativos de genes enriquecidos principalmente nas seguintes categorias de processos: proliferação, via, sinalização metabólica e dano ao DNA. As métricas do escore de enriquecimento, taxa de falsa descoberta e valores de P nominais foram utilizadas. Posteriormente, os DEGs foram enriquecidos nas vias metabólicas da base de dados do KEGG (*Kyoto Encyclopedia of Genes and Genomes*), e foram encontradas 12 vias enriquecidas que acrescentaram uma diversidade de funções biológicas, e três delas eram comuns a todas as três linhagens celulares de CG: proteólise mediada por ubiquitina, ribossomos, sistema e sinalização de células epiteliais em infecção por *Helicobacter pylori*. Neste estudo, as linhagens celulares de CG compartilham 14 genes regulados pelo *MYC*, mas o seu perfil de expressão gênica é diferente para cada subtipo histológico. Portanto, os resultados da análise *in silico* deste estudo revelaram as assinaturas de expressão relacionadas ao *MYC* no CG. Com isso, apresentamos evidências de que essas linhagens celulares de CG, representadas pelos subtipos histológicos distintos, têm diferentes perfis de expressão regulados pelo *MYC*, mas compartilham um núcleo comum de genes com perfis alterados. Esse é um passo importante para o entendimento do papel do *MYC* na carcinogênese gástrica; e também uma indicação de prováveis novos alvos de drogas em câncer de estômago.

Palavras-chave: NGS, siRNA, *MYC*, Linhagens Celulares, GSEA, KEGG, Câncer gástrico.

ABSTRACT

MYC is an oncogene responsible for excessive cell growth in cancer, allowing the transcriptional activation of genes involved in cell cycle regulation, metabolism and apoptosis, and is generally overexpressed in Gastric Cancer (GC). Using siRNA and *Next Generation Sequencing* (NGS), we identified the Genes Differential Expression (DEGs) regulated by *MYC* in three Brazilian cell lines of GC represented by the diffuse, intestinal and metastatic histological subtypes, and later integrated these data with a computational gene enrichment with the GSEA (*Gene Set Enrichment Analysis*) tool. We identified a total of 5,471 DEGs with a high correlation (80%). The silencing of *MYC* by siRNA in diffuse and metastatic CG cell lines resulted in an increase in the number of DEGs with decreased expression, while in intestinal-type lineage they exhibited a greater amount of DEGs with an increased expression profile. From gene enrichment, using our sequenced samples compared to the hallmark gene sets, we found 11 significant sets of genes enriched mainly in the following categories of processes: proliferation, pathway, metabolic signaling and DNA damage. Subsequently, DEGs were enriched in the metabolic pathways of the KEGG (*Kyoto Encyclopedia of Genes and Genomes*) database, and 12 enriched pathways were found that added a variety of biological functions, and three of them were common to all three cell lines of GC: ubiquitin-mediated proteolysis, ribosomes, system and epithelial cell signaling in *Helicobacter pylori* infection. In this study, GC cell lines shared 14 genes regulated by *MYC*, but their gene expression profile was different for each histological subtype. Therefore, the results of the *in silico* analysis of this study revealed expression signatures related to *MYC* in GC. Thus, we present evidence that these CG cell lines, represented by distinct histological subtypes, have different expression profiles regulated by *MYC*, but share a common nucleus of genes with altered profiles. This is an important step towards understanding the role of *MYC* in gastric carcinogenesis, as well as an indication of probable new drug targets in stomach cancer.

Key words: NGS, siRNA, *MYC*, Cell lines, GSEA, KEGG, Gastric Cancer.

LISTA DE ILUSTRAÇÕES

Figura 1. As características do câncer	13
Figura 2. Porcentagens de novos casos de câncer e mortes por câncer em todo o mundo em 2018	14
Figura 3. Distribuição proporcional dos dez tipos de câncer mais incidentes estimados para 2018 por sexo, exceto pele não melanoma, no Brasil	15
Figura 4. Distribuição proporcional dos dez tipos de câncer mais incidentes estimados para 2018 por sexo, exceto pele não melanoma, na região Norte do Brasil.....	16
Figura 5. Gráficos de vulcões de DEGs para três linhagens celulares de CG após o silenciamento do <i>MYC</i>	30
Figura 6. Painel com 11 conjuntos de genes enriquecidos para três linhagens de células de CG após o silenciamento do <i>MYC</i>	32
Figura 7. Escore de DEGs enriquecidos quando comparados com a coleção H com seu escore médio de enriquecimento (ES) para cada análise.....	34
Figura 8. Identificação de genes individuais nos conjuntos de dados DEG enriquecidos com o ES mais alto.....	36
Figura 9. DEGs comuns enriquecidos em três linhagens celulares de CG após o silenciamento <i>MYC</i>	39
Quadro 1. Diferenças entre os tipos difuso e intestinal de câncer gástrico.....	17

LISTA DE TABELAS

Tabela 1. Amostras utilizadas no estudo	25
Tabela 2. Leitura da quantificação após <i>MYC</i> -siRNA.....	29
Tabela 3. Conjuntos de características genéticas enriquecidos para DEGs após <i>MYC</i> -siRNA em três linhas celulares CG	31
Tabela 4. Os 14 DEGs comuns para as três linhagens de células de CG após o <i>MYC</i> -siRNA e as características de genes que foram enriquecidos em CG	38

APÊNDICES

APÊNDICE A. Artigo publicado	54
APÊNDICE B. Conjuntos de genes enriquecidos a partir da “KEGG” para DEGs	68
APÊNDICE C. 14 DEGs comuns para as três linhagens celulares de GC após o silenciamento do gene MYC-siRNA	69
APÊNDICE D. Os 14 DEGs comuns para as três linhagens celulares de GC após <i>MYC</i> -siRNA e as características dos genes que foram enriquecidos e identificados em estudos associados ao câncer	70

LISTA DE ABREVIATURAS

cDNA	DNA COMPLEMENTAR
CG	CÂNCER GÁSTRICO
DNA	ÁCIDO DESOXIRRIBONUCLEICO
ES	ENRICHMENT SCORE
FC	FOLD-CHANGE
FDR	FALSE DISCOVERY RATE
GEO	GENE EXPRESSION OMNIBUS
GSEA	GENE SET ENRICHMENT ANALYSIS
IARC	INTERNATIONAL AGENCY FOR RESEARCH ON CANCER
INCA	INSTITUTO NACIONAL DO CÂNCER
iRNA	RNA DE INTERFERÊNCIA
KEGG	KYOTO ENCYCLOPEDIA OF GENE AND GENOME
mRNA	RNA MENSAGEIRO
NES	NORMALIZED ENRICHMENT SCORE
NGS	NEXT GENERATION SEQUENCING
PCR	REAÇÃO EM CADEIA DA POLIMERASE
RNA-seq	SEQUENCIAMENTO DE RNA
ROS	REACTIVE OYGEN SPECIES
RPKM	READS PER KILOBASE PER MILLION MAPPED READS
RT-qPCR	PCR QUANTITATIVO DE TRANSCRIÇÃO REVERSA
siRNA	SHORT INTERFERING RNAS
TMAP	TORRENT MAPPING ALIGNMENT PROGRAM
WHO	WORLD HEALTH ORGANIZATION

SUMÁRIO

1	INTRODUÇÃO	12
1.1	Câncer	12
1.2	Câncer Gástrico	15
1.3	O oncogene <i>MYC</i> na carcinogênese gástrica	18
1.4	Cultura celular e linhagens gástricas	19
1.5	RNA de Interferência	20
1.6	Sequenciamento e enriquecimento usando GSEA	21
2	APLICABILIDADE CLÍNICA	23
3	OBJETIVOS	24
3.1	Objetivo Geral	24
3.2	Objetivo Específico	24
4	METODOLOGIA	25
4.1	Linhagens celulares e transfecção de siRNA	25
4.2	Sequenciamento de semicondutores e pré-tratamento de dados	25
4.3	Identificação e análise estatística de DEG's	26
4.4	Análise de enriquecimento funcional	27
4.5	PCR Quantitativo de Transcrição Reversa (RT-qPCR)	28
5	RESULTADOS	29
5.1	Quantificação de transcritos de CGs e Identificação de DEGs usando NGS	29
5.2	Linhagens Celulares de CG silenciadas por <i>MYC</i> de diferentes subtipos histológicos com perfis distintos de expressão dependentes de <i>MYC</i>	33
5.3	Os 14 DEGs regulados por <i>MYC</i> mostram perfis de expressão distintos para cada subtipo histológico de CG	40
6	DISCUSSÃO	41
7	CONCLUSÃO	44
8	REFERÊNCIAS	45
	APÊNDICE A – ARTIGO PUBLICADO	54
	APÊNDICE B	69
	APÊNDICE C	70
	APÊNDICE D	71

1 INTRODUÇÃO

1.1 Câncer

Os tecidos celulares são formados por conjuntos de células – alicerces dos organismos multicelulares – associados à matriz extracelular, conferindo a mecânica básica de sua estrutura. Seus componentes devem estar adequadamente organizados e coordenados, e muitos passam por manutenção e renovação constantes. Distúrbios nesses processos podem resultar na formação de tumores benignos ou no surgimento do câncer, caracterizado por ser a malignidade que está em primeiro lugar dentre as doenças de renovação dos tecidos e uma das principais causas de morte da humanidade (ALBERTS et al., 2017).

O Instituto Nacional do Câncer (INCA) denomina o câncer como o conjunto de mais de 100 doenças que apresentam em comum o crescimento desordenado de células com potencial de invadir tecidos e órgãos. Esse crescimento descontrolado pode ser causado, em alguns casos, por exposição a fatores ambientais (como dieta, luz solar ou componentes químicos); originando mutações em genes importantes (como o *MYC*) e assim, podem levar ao câncer. (SNUSTAD e SIMONS, 2013; ALBERTS, 2017; INCA, 2017). A acelerada divisão destas células tende a ser muito agressiva e incontrolável. Pode progredir para graus mais elevados de malignidade, transpassando a vasculatura, migrando para órgãos distantes do seu local de origem, atingindo tecidos saudáveis (livres de células cancerosas) e por fim podendo causar metástase à distância. Tais propriedades geralmente são ocasionadas por um acúmulo de mutações em classes críticas de genes, os oncogenes e os genes supressores de tumor, caracterizando o câncer como doença genética (FOSTER, 2008; LEAL et al., 2009).

As alterações acumuladas pelas células conferem vantagens que as diferenciam das demais, sendo selecionadas na evolução da tumorigênese e caracterizam a complexidade da biologia do câncer. Tais alterações foram classificadas e publicadas por Douglas Hanahan e Robert Allan Weinberg, no ano 2000, como sendo “As características do câncer” (*hallmarks*, em inglês) que formam um conjunto de capacidades funcionais adquiridas por alguns ou todos os tipos de câncer, mesmo que adquiridas por mecanismos variados. Os autores descreveram inicialmente seis características: (1) manutenção da sinalização proliferativa; (2) inibição da ação de supressores de crescimento; (3) resistência à morte celular programada; (4) potencial replicativo ilimitado; (5) ativação de invasão e metástase; (6) indução de angiogênese. Em

2011, os mesmos autores acrescentaram outras quatro *características*, totalizando as dez características principais para o desenvolvimento do câncer (Figura 1): (7) instabilidade genômica; (8) inflamação promovida pelo tumor; (9) reprogramação do metabolismo energético; (10) evasão do sistema imune (HANAHAN e WEINBERG, 2011; ROCHA, 2013; FOUAD e AANEI, 2017).



Figura 1. As características do câncer (Adaptado de HANAHAN e WEINBERG, 2011).

O câncer é diferenciado de acordo com o tipo de tecido inicialmente afetado, sendo conhecido como *carcinoma* quando tiver início em tecidos epiteliais como a pele ou mucosas, ou ainda como *sarcoma* quando iniciar em músculo ou em tecidos conjuntivos como osso e cartilagem (WEINBERG, 2008; INCA, 2018). Nas últimas décadas, significativos avanços foram obtidos acerca do seu entendimento, contudo ainda é considerado como grande problema de saúde pública em todo o mundo, sendo a segunda principal causa de morte em países desenvolvidos e a principal nos países em desenvolvimento (JEMAL et al., 2011).

São apontadas como prováveis causas do aumento geral no número absoluto de casos de câncer: a maior ocorrência de tipos de câncer relacionados com o pior prognóstico, o difícil acesso ao diagnóstico precoce e tratamento (UICC, 2018).

De acordo com a última estimativa mundial realizada pelo projeto

GLOBOCAN/IARC, divulgada este ano, haverá até o final de 2018 a somatória de 18,1 milhões de novos casos de câncer e 9,6 milhões de mortes em decorrência da doença, excentuando-se o câncer de pele não melanoma em ambas as situações. Tais estimativas baseiam-se em dados recentes disponíveis no IARC e em publicações disponíveis em plataformas virtuais. Os tipos de câncer com maiores índices de incidência e mortalidade no mundo em 2018 podem ser observados na figura 2 (IARC, 2018.1).

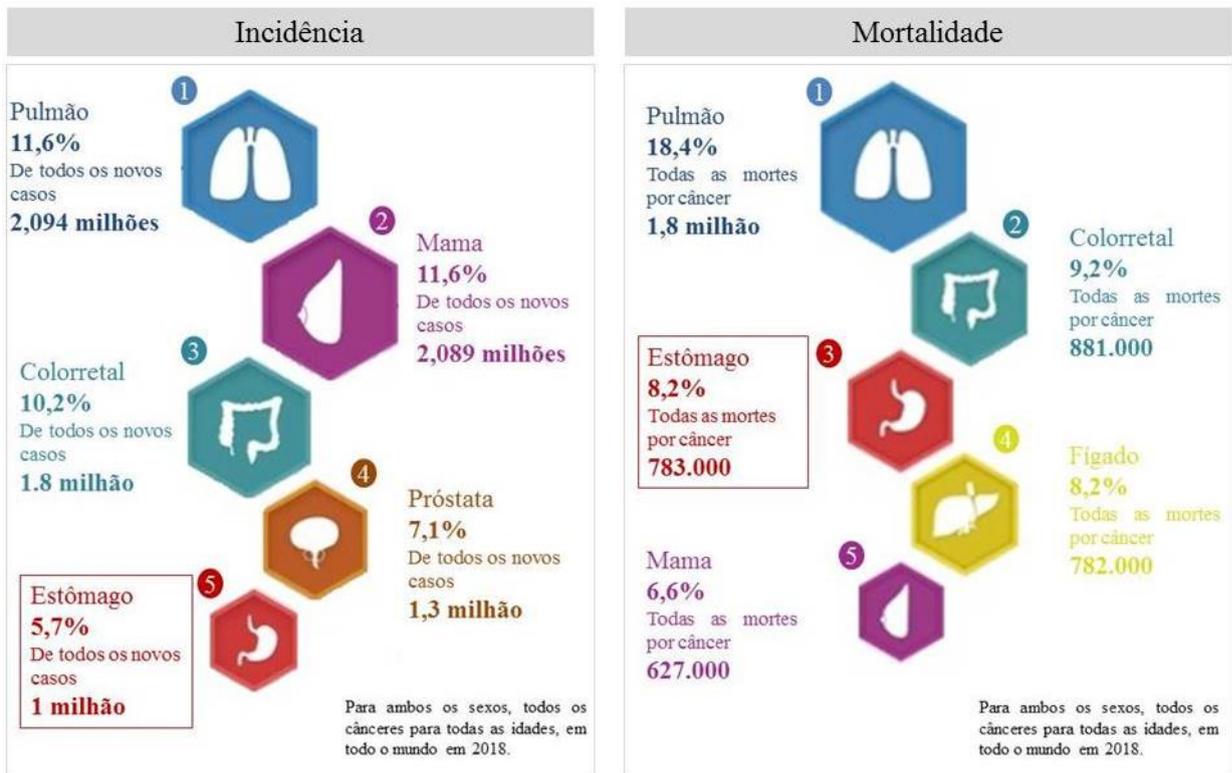


Figura 2. Porcentagens de novos casos de câncer e mortes por câncer em todo o mundo em 2018 (Adaptado de UICC, 2018).

Para o Brasil, a estimativa aponta a ocorrência de 600 mil novos casos de câncer em cada ano do biênio 2018 e 2019. Sendo aproximadamente 170 mil casos novos de câncer de pele não melanoma. Algumas variedades de câncer configuram um perfil de magnitude observados em países em desenvolvimento semelhante ao perfil de países desenvolvidos, principalmente com relação aos cânceres de próstata, mama e intestino. Contudo, ainda existe a relação de alguns tipos de câncer com condições socioeconômicas menos favoráveis, como o do colo do útero e o do estômago.

O Câncer Gástrico (CG) ocupa a quarta e a sexta posições dos mais incidentes em

homens e mulheres no Brasil, respectivamente (INCA, 2018). (Figura 3).

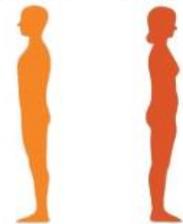
Localização primária	Casos	%			Localização primária	Casos	%
Próstata	68.220	31,7%		Homens Mulheres	Mama Feminina	59.700	29,5%
Traqueia, Brônquio e Pulmão	18.740	8,7%			Cólon e Reto	18.980	9,4%
Cólon e Reto	17.380	8,1%			Colo do Útero	16.370	8,1%
Estômago	13.540	6,3%			Traqueia, Brônquio e Pulmão	12.530	6,2%
Cavidade Oral	11.200	5,2%			Glândula Tireoide	8.040	4,0%
Esôfago	8.240	3,8%			Estômago	7.750	3,8%
Bexiga	6.690	3,1%			Corpo do Útero	6.600	3,3%
Laringe	6.390	3,0%			Ovário	6.150	3,0%
Leucemias	5.940	2,8%			Sistema Nervoso Central	5.510	2,7%
Sistema Nervoso Central	5.810	2,7%			Leucemias	4.860	2,4%

Figura 3. Distribuição proporcional dos dez tipos de câncer mais incidentes estimados para 2018 por sexo, exceto pele não melanoma*, no Brasil (INCA, 2018). * Números arredondados para múltiplos de 10.

Devido à gravidade da doença, métodos e técnicas de diagnóstico vêm sendo aperfeiçoados a fim de facilitar a detecção precoce, melhorar o tratamento e aumentar a sobrevivência global do paciente. Existem tratamentos que apresentam grandes chances de cura para pelo menos um terço dos casos de câncer, como o câncer de mama, colo do útero, cavidade oral e cólon, desde que detectados precocemente e tratados de forma adequada. De forma geral, são três as principais formas de tratamento do câncer: cirurgia, radioterapia e quimioterapia, que podem ser utilizadas em conjunto ou isoladas, dependendo do tipo e estadiamento do câncer (INCA, 2018). Mais recentemente vem sendo aplicada a terapia direcionada ou terapia molecular alvo que utiliza inibidores químicos que interagem bloqueando moléculas específicas do câncer e vias de sinalização, e é considerado um tratamento promissor para vários tipos de câncer, inclusive para o CG (NAM et al., 2013).

1.2 Câncer Gástrico

O câncer gástrico (CID-10 C16) é uma das principais causas de morte relacionada ao câncer no mundo, mesmo com o declínio da sua incidência na maioria dos países desenvolvidos ao longo das últimas três décadas. No Brasil é um dos mais incidentes em ambos os sexos. Para cada ano do biênio 2018/2019 é esperada a ocorrência de 13.540 novos casos de câncer de estômago entre homens e 7.750 nas mulheres (INCA, 2018).

Na região Norte do país, o câncer gástrico é o segundo mais frequente entre os homens (12,35/100 mil) e o quinto mais frequente entre as mulheres (5,34/100 mil), desconsiderando o câncer de pele não melanoma (Figura 4). O estado do Pará apresenta taxas de mortalidade historicamente acima da média geral dos demais estados (GUIMARÃES e MUZI, 2012; INCA, 2018).

Localização primária	Casos	%			Localização primária	Casos	%
Próstata	2.700	28,8%	Homens		Colo do Útero	2.300	24,8%
Estômago	1.150	12,2%			Mama Feminina	1.730	18,6%
Traqueia, Brônquio e Pulmão	820	8,7%			Cólon e Reto	660	7,1%
Cólon e Reto	450	4,8%			Traqueia, Brônquio e Pulmão	520	5,6%
Leucemias	390	4,2%			Estômago	480	5,2%
Cavidade Oral	330	3,5%			Leucemias	310	3,3%
Linfoma não Hodgkin	270	2,9%			Ovário	270	2,9%
Sistema Nervoso Central	270	2,9%			Glândula Tireoide	270	2,9%
Laringe	240	2,6%			Sistema Nervoso Central	270	2,9%
Esôfago	240	2,6%			Corpo do Útero	200	2,2%

Figura 4. Distribuição proporcional dos dez tipos de câncer mais incidentes estimados para 2018 por sexo, exceto pele não melanoma*, na região Norte do Brasil (INCA, 2018). * Números arredondados para múltiplos de 10.

Fatores ambientais são importantes fatores de risco para o câncer gástrico. A infecção pela bactéria *Helicobacter pylori* (*H. pylori*), classificada como carcinogênica do grupo I é um desses principais agentes, liderando as infecções que causam câncer (IARC, 2018). Outras causas são: o consumo de alimentos conservados em sal e defumados, o baixo uso de refrigeradores e o baixo consumo de frutas e vegetais frescos (NAKASHIMA et al., 2012). Também são apontados como fatores de risco associados, o tabagismo, histórico familiar de CG, anemia perniciosa, modificações genéticas e/ ou epigenéticas e o grupo sanguíneo A (BERMÚDEZ et al., 2006; ZILBERSTEIN, 2012; YANG et al., 2014).

A carcinogênese gástrica é altamente heterogênea com variadas características genéticas e moleculares podendo se originar de diferentes formas histológicas, como por exemplo: *adenocarcinoma* (originado em células glandulares, ocorre na maioria dos casos de câncer gástrico), *linfoma* (relacionados ao sistema linfático) e *leiomiosarcoma* (originado em células do musculo liso) (INCA, 2017).

Os adenocarcinomas podem ser divididos em dois tipos patológicos distintos: *intestinal* e *difuso*, com diferentes características e prognósticos (LAURÉN, 1965). O tipo intestinal apresenta maior número de cópias do *MYC*, enquanto o tipo difuso apresenta maior expressão deste gene e mutações pontuais mais frequentemente do que em tumores do tipo intestinal, sugerindo tratar-se de duas entidades distintas (DE SOUZA, 2013).

O CG do tipo difuso é considerado *de novo*, pois não evolui a partir de lesões pré-cancerosas e células precursoras ainda não foram identificadas, e está associado a pacientes mais jovens, com maior ocorrência no sexo feminino. Neste tipo de câncer as células tumorais não se apresentam aderidas entre si, devido à perda de moléculas de adesão, como a E-caderina, uma das principais proteínas de superfície celular no estabelecimento de conexões intercelulares e manutenção da organização dos tecidos epiteliais além de estar relacionada a uma rápida progressão, ocorrência de metástases peritoneais, pior prognóstico e afetar principalmente o corpo do estômago. Já o câncer gástrico do tipo intestinal é, geralmente, associado à metaplasias intestinais, neoplasia intraepitelial e carcinoma. As células tumorais apresentam maior adesão entre si e os tumores possuem progressão lenta, acometem comumente pacientes mais velhos e do sexo masculino, estão amplamente associados à invasão linfática ou vascular, apresentam melhor prognóstico e afetam principalmente a região antro gástrica do estômago (YAMASHITA et al., 2011; MA et al., 2016) (Quadro 1).

TIPO INTESTINAL	TIPO DIFUSO
Cópias de <i>MYC</i> elevadas	Hipometilação do <i>MYC</i>
Progressão lenta	Rápida progressão / Metástases peritoneais
Pacientes mais velhos/ sexo masculino	Pacientes jovens/ sexo feminino
Melhor prognóstico	Pior prognóstico
Afeta a região antro gástrica do estômago	Afeta o corpo do estômago

Quadro 1. Diferenças entre os tipos difuso e intestinal de câncer gástrico.

Inúmeros fármacos citotóxicos, esquemas terapêuticos e recentes avanços em técnicas cirúrgicas combinadas com quimioterapia neoadjuvante e abordagens de radioterapia estão disponíveis para o tratamento do CG, no entanto, os pacientes com CG avançado ainda têm resultados pouco expressivos (VAN NESS et al., 2012). O tratamento mais efetivo para o câncer gástrico é a gastrectomia (remoção completa do estômago) uma vez que a doença é detectada mais frequentemente em estágio avançado e a quimioterapia e radioterapia não são mais eficazes (ZALI, REZAEI-TAVIRANI e AZODI, 2011; SILVA, 2017).

1.3 O Oncogene *MYC* na Carcinogênese Gástrica

O gene *MYC* apresenta três éxons e é mapeado na região cromossômica 8q24.21 (braço longo do cromossomo 8, região 2, banda 4, sub bandas 2 e 1) (NCBI, 2018). Desempenha papel importante em funções fundamentais da biologia celular, dentre elas, a regulação do crescimento e proliferação celular, metabolismo, diferenciação, apoptose e angiogênese. Alterações em qualquer uma dessas funções pode contribuir para o fenótipo do tumor. Apresenta efeito em diversos genes dos genomas de muitos organismos, incluindo o de humanos, e é amplamente estudado devido à sua associação com várias doenças. (DE SOUZA, 2013; FERNANDEZ, 2003).

Mudanças nesse gene estão geralmente associadas à amplificação gênica, translocação ou inserção cromossômica, mutações, resistência da proteína *MYC* à proteólise mediada pela ubiquitina e modificações epigenéticas, sendo frequentemente observadas em diferentes tipos de câncer, incluindo o gástrico. O *MYC* se expressa elevada e desreguladamente em muitos casos de neoplasias humanas, e acredita-se estar relacionado diretamente a vias e processos importantes na carcinogênese. Estudos demonstraram que a expressão do seu mRNA, bem como o número de cópias, aumentam durante o processo da carcinogênese gástrica do tipo intestinal, relacionando o *MYC* à iniciação e progressão do tumor gástrico (DE SOUZA, 2013). Este gene também se encontra desregulado em lesões pré-neoplásicas gástricas, podendo ter um papel na carcinogênese gástrica precoce (CALCAGNO et al., 2008).

Calcagno e colaboradores (2006) conduziram um estudo com sete casos de adenocarcinoma gástrico primário (5 do tipo difuso e 2 do tipo intestinal) submetidos à ressecção, utilizando as técnicas de imunohistoquímica e FISH, que apresentou o *MYC* expresso em todas as amostras, com amplificação elevada exclusivamente no tipo intestinal e imunorreatividade deste gene apenas no núcleo, em contrapartida, rearranjos estruturais, como translocação, foram observados somente no tipo difuso, com imunorreatividade no núcleo e no citoplasma. Esses padrões distintos de alterações entre os dois tipos histológicos reforçam a hipótese de que eles são guiados por diferentes mecanismos de tumorigênese genética.

Níveis excessivos de *MYC* intracelular causam instabilidade genômica (WISNIESKI, 2013) e replicação de DNA propensa a erros causados por estresse replicativo induzido por oncogene (WHITFIELD e SOUCEK, 2011). Também pode ser um indicativo do alto grau de agressividade do tumor uma vez que o número de cópias do gene está diretamente relacionado ao seu nível de expressão (CALCAGNO, 2006). A identificação de genes regulados por *MYC*

geralmente baseia-se na ativação experimental de *MYC*, seguida de monitoramento de alterações nos níveis de mRNA (GRANDORI et al., 2000; OSTER et al., 2002).

Embora já se saiba que o câncer gástrico pode estar associado ao aumento da expressão de *MYC*, muitos estudos genéticos de alto rendimento não evidenciam sua importância na tumorigênese gástrica (ZHANG et al., 2010; WALZ e EILERS, 2013; HILLS e DIFFLEY, 2014; MCLEAN, M. H. e EL-OMAR, 2014; FIGUEIREDO et al., 2017). Dessa maneira, a inibição da expressão desse gene pode ser uma ferramenta potencial para a elucidação dos processos envolvidos no CG e também para o tratamento deste (VITA e HENRIKSSON, 2006).

1.4 Cultura Celular e Linhagens Gástricas

O cultivo de células *in vitro* vem sendo estudado desde o início do século passado como um método para compreender o comportamento de células animais fora do organismo, em um ambiente controlado com nutrientes e temperatura adequados que mimetizam as condições naturais do seu local de origem (ROCHA, 2013).

Foi primeiramente desenvolvido por Ross Granville Harrison, em 1907, quando estudou o desenvolvimento das fibras nervosas; tomando por base esse estudo, Carrel, em 1912, desenvolveu um modelo para cultivo a partir de células cardíacas de embrião de galinha, onde observou a necessidade da troca periódica dos nutrientes fornecidos à cultura. Estudos posteriores ajudaram a definir a metodologia conforme é desenvolvida até hoje, e desde então é amplamente utilizada nas pesquisas realizadas em laboratórios do mundo inteiro (HARRISON, 1907).

A linhagem celular CaVe descrita por Dobrynin (1963) foi a primeira a ser estudada a partir do câncer gástrico. Poucas linhagens celulares de tumores gástricos estão disponíveis apesar da alta incidência da doença e, aproximadamente 80% delas foram desenvolvidas em países asiáticos (KU et al., 2012), onde epidemiologicamente a sua prevalência é maior, principalmente em associação com infecção por *H. pylori*. No Japão, o câncer gástrico é um dos mais frequentes tipos de câncer em ambos os sexos e acomete cerca de 20 a 30% de toda a incidência por câncer, o que justifica esse interesse no estudo com linhagens de câncer gástrico. (SUZUKI e SEKIGUCHI, 2002; LIMA et al., 2004).

Estudos com linhagens celulares de câncer humano têm gerado importantes

esclarecimentos a respeito da biologia das células neoplásicas, processos patológicos e o desenvolvimento de novas estratégias contra o crescimento e a progressão dessas células (FRESHNEY, 2010; BARRETINA et al., 2012), visando o aumento da sobrevida global e a redução da taxa de mortalidade com a promoção de novos manejos do paciente e estratégias terapêuticas (MELLO et al., 2015).

As linhagens de CG utilizadas no presente estudo foram estabelecidas e caracterizadas pelo grupo de pesquisa do Laboratório de Citogenética Humana da Universidade Federal do Pará: AGP01 (Ascite Gástrica Paraense 01), ACP02 e ACP03 (Adenocarcinoma Gástrico Paraense 02 e 03, respectivamente) a partir de amostras de pacientes oriundos do estado do Pará. Análises cariotípicas apresentaram a deleção do braço curto do cromossomo 17 (17p) e a trissomia dos cromossomos 7 e 8 como as alterações cromossômicas mais comuns às três linhagens; amplificação de *MYC* e deleção do gene *TP53*, são alterações genéticas comuns em pacientes com câncer gástrico brasileiro (KHAYAT, 2009). As linhagens AGP01 e ACP03 apresentam em comum a trissomia do cromossomo 17 (LEAL et al., 2009).

A linhagem AGP01 foi estabelecida a partir de células neoplásicas metastáticas obtidas de um aspirado de líquido ascítico de um paciente do sexo masculino com adenocarcinoma gástrico do tipo intestinal, em estágio T3N2M1. A linhagem ACP02 foi isolada de adenocarcinoma gástrico primário, a partir de um tumor do tipo difuso em estágio T3N2M0. A linhagem ACP03 foi estabelecida a partir de tumor gástrico primário do tipo intestinal em estágio mais avançado (T4N1M0) (LIMA et al., 2004; LEAL et al., 2009).

1.5 RNA de Interferência (iRNA)

O processo de interferência mediada por RNA é um mecanismo natural evolutivamente conservado nas células de plantas e animais, cuja principal função é proteger tais células do ataque de determinados vírus e transposons, bem como manter a estabilidade do material genético. Os genomas destes invasores são compostos por duas fitas de RNA. Ao adentrar no citoplasma da célula é reconhecido como um corpo estranho e há a ativação da enzima nuclease *Dicer* que cliva o genoma viral em fragmentos de 21-25 nucleotídeos, inativando-o dentro da célula (BARBOSA e LIN, 2004).

siRNA é uma técnica com ampla difusão nos laboratórios de genética molecular

e representa um ícone da genética moderna. Baseia-se na interferência da expressão de genes nas células sem alterar seu material genético (DNA), mediado por pequenos RNAs de fita dupla (dsRNA- *double stranded RNA*) que reconhecem especificamente uma sequência de RNA mensageiro (mRNA) alvo resultando na sua clivagem ou repressão traducional. Utilizada em diversas aplicações como em doenças cardiovasculares e reumáticas, desordens genéticas herdadas, infecções virais e câncer (PEREIRA, 2013).

Esta técnica foi utilizada no presente trabalho para bloquear a tradução do mRNA do gene *MYC*, seguido pelo sequenciamento íon proton (sequenciamento do semicondutor de íons), a fim de identificar a assinatura de regulação do *MYC* nas linhagens celulares AGP01, ACP02 e ACP03 comparando amostras silenciadas e não silenciadas.

1.6 Sequenciamento e Enriquecimento Usando GSEA

A primeira técnica de sequenciamento foi desenvolvida pelo bioquímico inglês Frederick Sanger, em 1977 (IBQM, 2018) e foi utilizada para decifrar o genoma humano. Com o avanço das pesquisas, os equipamentos foram assumindo formas mais refinadas permitindo maior sensibilidade e precisão nos resultados das metodologias aplicadas. O Sequenciamento de Nova Geração (NGS - *Next Generation Sequencing*, em inglês), é um conjunto de técnicas que surgiu no ano de 2005 que captura um espectro mais amplo de mutações comparado à técnica de Sanger (BEHJATI e TARPEY, 2013).

NGS é alimentado por chips semicondutores que ajudam a implementar um fluxo de trabalho rápido e simples que atende às necessidades de pesquisa em várias aplicações, como na oncologia. Com o seu advento é possível estudar sistematicamente genomas do câncer agora na íntegra e o sequenciamento individual do câncer pode fornecer a base do seu tratamento personalizado (THERMO FISHER, 2018.1).

A metodologia baseia-se na fragmentação do DNA em pequenos segmentos que posteriormente recebem adaptadores que possibilitam o sequenciamento do fragmento. A inserção dos nucleótidos específicos pode ser medida por luz (Illumina, San Diego, CA) ou por alterações de pH (Ion Torrent; ThermoFisher Scientific, Nova Iorque, NY). Resultando na geração de milhões de segmentos medindo de 50 a 250 pares de bases (pb) que serão comparados a um genoma ou mapa de referência. A análise de dados de bioinformática do conjunto de dados gerado é processada por computadores de alto desempenho. O RNA pode

ser sequenciado convertido em DNA complementar (cDNA) usando a enzima transcriptase reversa, seguido do mesmo procedimento aplicado para o DNA (ROYCHOWDHURY e CHINNAIYAN, 2016).

A bioinformática tem sido aplicada principalmente em pesquisas científicas básicas. Após a conclusão do sequenciamento do genoma humano, ela facilitou inúmeras descobertas em medicina básica, e várias aplicações clínicas de bioinformática foram relatadas, incluindo o sequenciamento clínico, um campo emergente de medicina de precisão (KATO, 2016). Na pesquisa do câncer, a bioinformática tem sido usada para estudar o transcriptoma do câncer, o diagnóstico precoce, a classificação e a previsão do prognóstico (SUN et al., 2017).

Atualmente, uma nova era vem ganhando espaço na pesquisa, a “era das ômicas” que se configura na coleta e análise de conjuntos de dados de variáveis biológicas em larga escala. Regularmente aplicada a regiões codificantes de genes (genômica), RNAs mensageiros e *splicing* alternativo (transcriptômica), proteínas (proteômica) e suas subunidades, incluindo nucleotídeos e aminoácidos (metabolômica) (ROYCHOWDHURY e CHINNAIYAN, 2016).

Neste estudo, usamos NGS baseado em semicondutores, para quantificar os transcritos e suas isoformas nas linhagens AGP01, ACP02 e ACP03 antes e depois do silenciamento do *MYC* utilizando siRNA.

2 APLICABILIDADE CLÍNICA

O gene *MYC* é altamente importante na carcinogênese gástrica, ativando e inativando genes nas várias etapas da doença. O câncer Gástrico apresenta ampla heterogeneidade, a caracterização genética dos diferentes tipos desta doença pode auxiliar no diagnóstico, além de enriquecer a base de dados sobre os diferentes subtipos histológicos deste câncer.

A identificação dos genes que tem sua expressão influenciada por *MYC* também é importante para a definição de possíveis novos alvos terapêuticos. Possibilidades de diagnósticos e terapias mais específicas para cada tipo da doença são importantes alternativas para auxiliar no entendimento desse câncer que ainda é um dos cânceres de pior prognóstico.

3 OBJETIVOS

3.1 Objetivo Geral

Identificar o perfil da expressão diferencial de genes (DEGs) em três linhagens celulares de câncer gástrico (AGP01, ACP02 e ACP03) após o silenciamento do *MYC*.

3.2 Objetivos Específicos

1. Quantificar os DEGs nas amostras das três linhagens celulares de CG sequenciadas com Ion Proton após o silenciamento do *MYC*;

2. Integrar os resultados dos DEGs de todas as amostras das três linhagens de CG com a ferramenta computacional GSEA (Gene Set Enrichment Analysis), para identificar os genes enriquecidos como prováveis alvos regulados pelo *MYC*;

3. Identificar as vias metabólicas de interesse associadas com prováveis alvos regulados pelo *MYC*;

4 METODOLOGIA

4.1 Linhagens Celulares e Transfecção de siRNA

Foram usadas as linhagens celulares de CG: AGP01 (metastática), ACP02 (tipo difuso) e ACP03 (tipo intestinal). A linhagem MN01, uma linhagem de células da mucosa gástrica não neoplásica também foi utilizada para avaliar a expressão gênica e proteica após o silenciamento do *MYC*, comparando esses resultados com os obtidos com as linhagens de CG.

Foram semeadas 300 mil (3×10^5) células em placas de cultura de 6 cm² de cada linhagem celular durante 24h antes da transfecção. Pequenos RNAs de interferência direcionados à *MYC* ou siRNAs de controle foram transfectados nas linhagens celulares AGP01, ACP02 e ACP03 utilizando Reagente de Transfecção de Lipofectamina RNAiMAX. A transfecção ótima foi atingida após 48h, e o RNA total e proteínas foram extraídos com reagente TRIzol. Os reagentes foram adquiridos comercialmente de Thermo Fisher Scientific, EUA. Todos os experimentos com siRNA foram realizados três vezes. A Tabela 1 apresenta os nomes das amostras e os códigos de acesso Gene Expression Omnibus (GEO).

Tabela 1. Amostras utilizadas no estudo.

Linhagem Celular	Amostras (GEO)	Nome da Amostra	Subtipo histológico CG
AGP01	GSM2147866	1C	Flúido ascítico de CG
	GSM2147867	1M	Flúido ascítico de CG
ACP02	GSM2147868	2C	Difuso
	GSM2147869	2M	Difuso
ACP03	GSM2147870	3C	Intestinal
	GSM2147871	3M	Intestinal

As amostras do controle de si-RNA são rotuladas como *C* e *MYC*-siRNA como *M*.

4.2 Sequenciamento de Semicondutores e Pré-tratamento de Dados

Amostras de RNA total foram primeiramente tratadas com DNase-I para remover qualquer possível contaminação com DNA, em seguida o mRNA foi enriquecido utilizando

Dynabeads Oligo (dt) 25 obtida comercialmente da Thermo Fisher Scientific, USA. O mRNA enriquecido foi clivado em fragmentos menores de aproximadamente 200 pares de bases, que foram anexados a adaptadores com sequências conhecidas, únicas para cada amostra. As amostras foram conectadas a esferas magnéticas contendo sequências complementares aos adaptadores e então inseridas em micropoços onde foi realizada a PCR em emulsão para a síntese de cDNA (illustra Ready-To-Go RT-PCR Beads, GE Lifesciences). As seis bibliotecas de cDNA foram submetidas à quantificação e controle de qualidade utilizando o Agilent 2100 Bioanalyzer e então foram carregadas em chip Ion Proton V2 PI usando o Kit de Sequenciamento Ion PI 200 v3 e sequenciados utilizando o sequenciamento do semicondutor de íons (Thermo Fisher Scientific, EUA) em uma única execução multiplex.

As leituras de dados brutos que foram obtidas por sequenciamento primário usando Ion Proton foram submetidas ao controle de qualidade para calcular o alinhamento e avaliar como as leituras se comportam quando comparadas ao genoma humano de referência (Hg19/GRCh37). As leituras alinhadas foram mapeadas e quantificadas utilizando o Programa de Alinhamento de Mapeamento Torrent (TMAP - Torrent Mapping Alignment Program), que suporta diferentes algoritmos de alinhamento (LI e DURBIN, 2010; LENG et al., 2013; SUN et al., 2017). Os conjuntos de dados processados foram carregados no GEO com o número de acesso GSE81265.

4.3 Identificação e Análise Estatística de DEGs

O método computacional (*in silico*) GSEA utiliza abordagens estatísticas para identificar a ocorrência de diferenças estatisticamente significantes entre dois estados biológicos em classes de genes ou proteínas enriquecidas que se apresentaram em conjuntos que possivelmente têm associação com características fenotípicas de doenças. O Banco de Dados de Assinatura Molecular (MSigDB - Molecular signatures database, em inglês) (LIBERZON et al., 2015) divide os conjuntos de genes em oito coleções, dentre as quais, duas foram utilizadas no presente trabalho: a coleção H (grupo de genes Hallmark) que apresenta “assinatura coerentemente expressa derivada pela agregação de muitos conjuntos de genes MSigDB para representar estados ou processos biológicos bem definidos” (SUBRAMANIAN et al., 2005) e a coleção C2 (Conjunto de genes curados: KEGG) a partir de bases de dados disponíveis na plataforma disponível publicamente que é alimentada por fontes como as

publicações no PubMed e conhecimentos de especialistas (BROAD INSTITUTE, 2018).

A análise de Enriquecimento de Conjuntos de Genes (BENJAMINI e YEKUTIELI, 2001) foi utilizada para identificar conjuntos de genes significativamente enriquecidos entre o controle de siRNA e o *MYC* silenciado e vice-versa.

O pacote de software Sailfish (LI, 2012) e as RPKM - Leituras por Quilobase por Milhão de Leituras Mapeadas (Reads Per Kilobase per Million Mapped Reads, em inglês) (TRAPNELL et al., 2010; PATRO, MOUNT e KINGSFORD, 2014) foram usadas para analisar a significância de DEGs entre amostras do controle e *MYC* silenciado. Para identificar os DEGs entre as amostras pareadas foi utilizado o teste Audic-Claverie (GARBER, 2011). O Fold-Change (FC) foi calculado como a razão Log_2 entre a amostra silenciada (M) e a amostra controle (C). Foi usada uma correção de p-valor correspondente a testes de expressão diferencial usando correção Bonferroni (AUDIC e CLAVERIE, 1997). Nosso corte para definição de DEGs foi estabelecido como uma Taxa de Falsa Descoberta (FDR) < 0.05 (ABDI, 2007) e $|\text{Log}_2(\text{FC})| > 1$. DEGs foram plotados usando o Multiplot v2, que exibe perfis de expressão gênica personalizada.

4.4 Análise de Enriquecimento Funcional

As alterações de expressão genética observadas por DEGs foram relacionadas com vias biologicamente enriquecidas encontradas em GSEA. Os conjuntos de dados de expressão gênica utilizados foram coleções H (grupo de genes Hallmark) (SUBRAMANIAN et al., 2005) e C2 (conjunto de genes curados: KEGG), publicamente disponíveis no Banco de Dados de Assinatura Molecular (MsigDB - Molecular signatures database, em inglês) (LIBERZON et al., 2015). Os parâmetros padrão definidos por Subramanian e colaboradores foram utilizados em nossa análise. A significância estatística da análise GSEA foi determinada por 100 permutações, os mapas de enriquecimento foram criados para conjuntos de genes significantes ($P < 0.05$ e Taxa de Falsa Descoberta (FDR) < 0.25) e o Software GraphPad Prism foi utilizado para representar graficamente nossos dados.

4.5 PCR Quantitativo de Transcrição Reversa (RT-qPCR)

Para confirmar o efeito silenciador do siRNA na expressão de *MYC*, utilizamos a PCR quantitativa em tempo real para avaliar sua expressão em relação à expressão encontrada para MN01. Todos os testes foram feitos em triplicata e usando sondas TaqMan como produtos de testes sob demanda para expressão gênica (Life Technologies, EUA) (*MYC*: Hs00153408_m1). Os 14 DEGs identificados (ANEXO C) também foram analisados e validados por PCR quantitativa em tempo real. A expressão desses genes foi calculada em relação a sua expressão na MN01.

5 RESULTADOS

5.1 Quantificação de Transcritos de CGs e Identificação de DEGs Usando NGS

O sequenciamento NGS de seis bibliotecas gerou mais de 75 milhões de leituras, as quais, após o enriquecimento, foram mapeadas dentro do genoma de referência em mais de 99% das amostras e em mais de 98% do transcriptoma de referência (Tabela 2), e a distribuição dos segmentos amplificados foi consistente em todas as amostras. A tabela 2 também mostra em média quantos genes foram identificados para cada amostra. As leituras médias produzidas pelo sequenciamento do semicondutor de íons variaram entre 125 e 130 pbs.

Tabela 2. Leitura da quantificação do *MYC* após o silenciamento por siRNA.

Nome da Amostra	Total de Leituras	Total de Leituras Mapeadas	Taxa de Mapeamento Genético	Genes Expressos
1C	13.319.736	12.944.452	97.18%	14.594
1M	11.043.607	10.771.298	97.53%	10.709
2C	12.021.142	11.751.830	97.76%	9.859
2M	11.701.695	11.481.863	98.12%	8.793
3C	12.772.535	12.546.479	98.23%	7.988
3M	12.430.807	12.112.157	97.44%	10.685

De acordo com o ponto de corte ($FDR < 0.05$ e $|\text{Log}_2(FC)| > 1$), obteve-se uma quantidade distinta de DEGs entre o controle de siRNA e amostras de *MYC* silenciadas. Usando Multiplot (v2) foram identificados 1.556 DEGs com expressão diminuída e 917 com expressão aumentada na linhagem AGP01; foram encontrados 4.098 DEGs com expressão diminuída e 1.229 com expressão aumentada na linhagem celular do tipo difuso (ACP02); e, identificamos 3.272 DEGs com expressão aumentada e 842 com expressão diminuída para a linhagem ACP03 (Figura 5). Nossos resultados indicam que é possível discernir os subtipos histológicos de CG analisando o padrão de expressão gênica relacionado ao *MYC*.

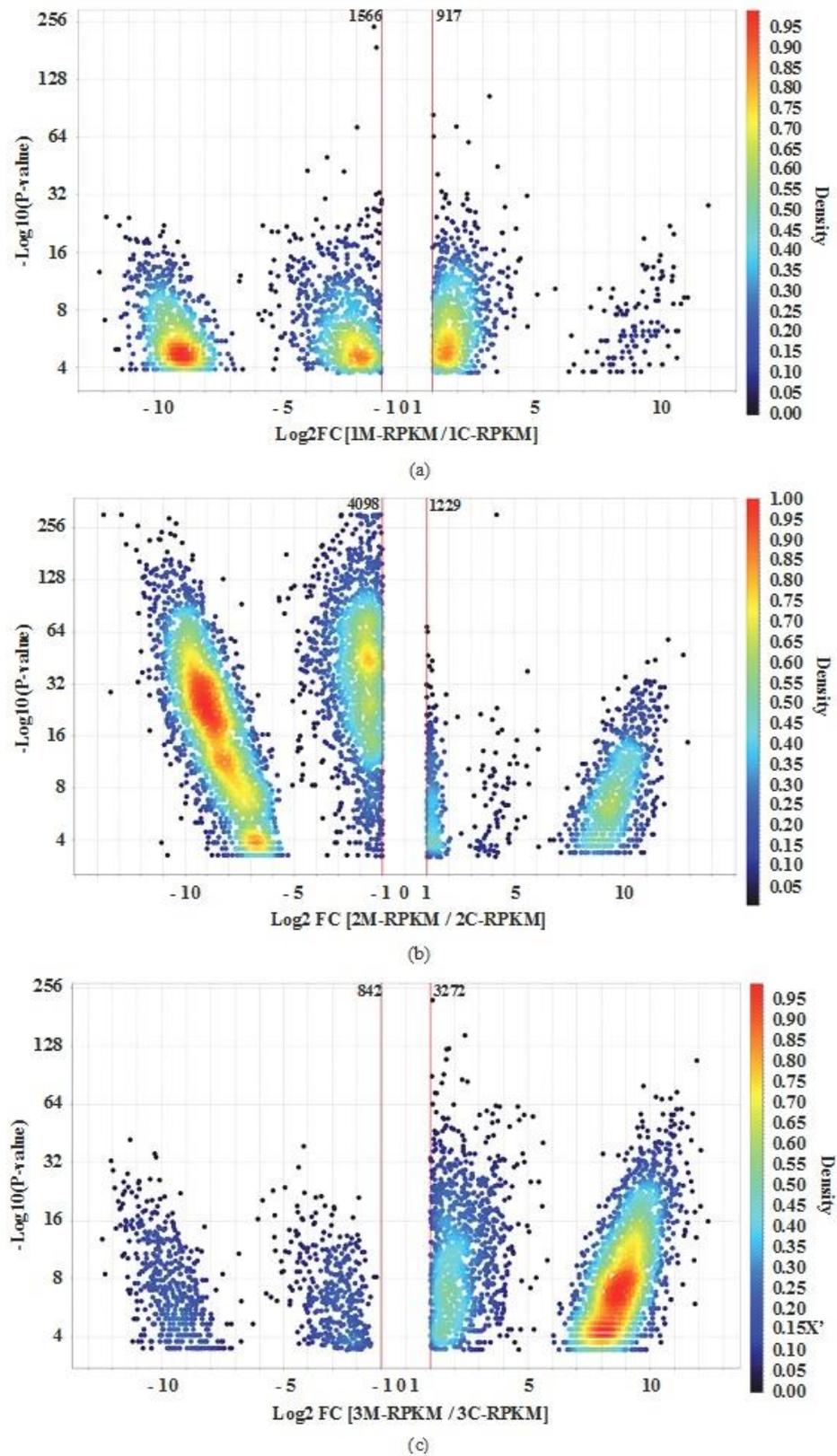


Figura 5. Gráficos de vulcões de DEGs para três linhagens celulares de CG após o silenciamento do *MYC*. Log₂ Fold-change e valores-p ($-\text{Log}_{10}$) são mostrados para DEGs com ($|\text{Log}_2(\text{FC})| > 1$ e $p \leq 0,05$). (A) DEGs para AGP01. (B) DEGs para ACP02. (C) DEGs para ACP03. Densidade é um cálculo de caso especial que é usado apenas em sombreamento avançado. Considera apenas que os operandos para este cálculo são os valores X e Y dos eixos do gráfico.

Um total de 16.777 genes dos seis conjuntos de dados obtidos por RNA-seq (GSE81265) foi inserido em uma matriz de expressão normalizada por RPKM e, após isso, usado para o enriquecimento comparando com grupos de dados de expressão das coleções H e C2. A aplicação do GSEA objetivou procurar conjuntos de genes que apresentavam enriquecimento apenas em amostras de controle de siRNA, mas não em amostras de *MYC* silenciadas, provavelmente alvos de *MYC* e foram encontrados 11 conjuntos de genes significantes ($P < 0.05$ e $FDR < 0.25$), como mostrado na Tabela 3.

Tabela 3. Conjuntos de características genéticas enriquecidos para DEGs após *MYC*-siRNA em três linhas celulares CG.

Nome da característica:	Categoria de função	Descrição	Tamanho*	ES	NES	Nom. P-valor	FDR	Max. RANK
ALVOS <i>MYC</i> V1	Proliferação	Alvos do <i>MYC</i> , variante 1	100/44	0.59	1.89	0.000	0.000	1368
VIA DE ESPÉCIES REATIVAS DE OXIGÊNIO	Via	Via de espécies reativas de oxigênio	25/13	0.55	1.42	0.000	0.000	2316
RESPOSTA A PROTEÍNAS MAL ENOVELADAS	Via	Resposta proteica desdobrada; Estresse ER	50/22	0.55	1.56	0.000	0.000	1891
SECREÇÃO PROTEICA	Via	Secreção de proteínas	44/26	0.55	1.70	0.000	0.000	2376
G2M_CHECKPOINT	Proliferação	Progressão do ciclo celular: ponto de verificação G2 / M	97/40	0.52	1.67	0.000	0.000	1813
SINALIZAÇÃO MTORC1	Sinalização	Sinalização mTORC1	102/43	0.50	1.46	0.000	0.044	2021
ALVOS E2F	Proliferação	Progressão do ciclo celular: alvos E2F	89/35	0.48	1.42	0.000	0.052	1607
FOSFORILAÇÃO OXIDATIVA	Metabólico	Ciclo de fosforilação oxidativa e ácido cítrico	118/57	0.48	1.46	0.000	0.062	2284
HOMEOSTASE DE COLESTEROL	Metabólico	Homeostase do colesterol	39/14	0.48	1.48	0.000	0.069	1671
REPARO DE DNA	Danos no DNA	Reparo de DNA	80/38	0.46	1.42	0.000	0.077	2337
APOPTOSE	Via	Morte celular programada; Via da caspase	66/29	0.42	1.31	0.200	0.190	2405

*Número de genes encontrados para cada conjunto de genes após a análise de dados. ES: escore de enriquecimento. NES: escore de enriquecimento normalizado. Nom. Valor P: valores P nominais para cada conjunto de genes. FDR: taxa de falsa descoberta. Max. rank: números máximos de classificação para cada conjunto de genes.

Foi encontrado um total de 7.903 genes, e 5471 (69,2%) enriquecidos em amostras de controle de siRNA, apresentando uma correlação muito alta entre réplicas biológicas e bibliotecas (79,1%). Os mapas de enriquecimento obtidos são mostrados na Figura 6, onde criamos um painel de 11 conjuntos de genes significativamente enriquecidos.

Enplots

Heatmaps

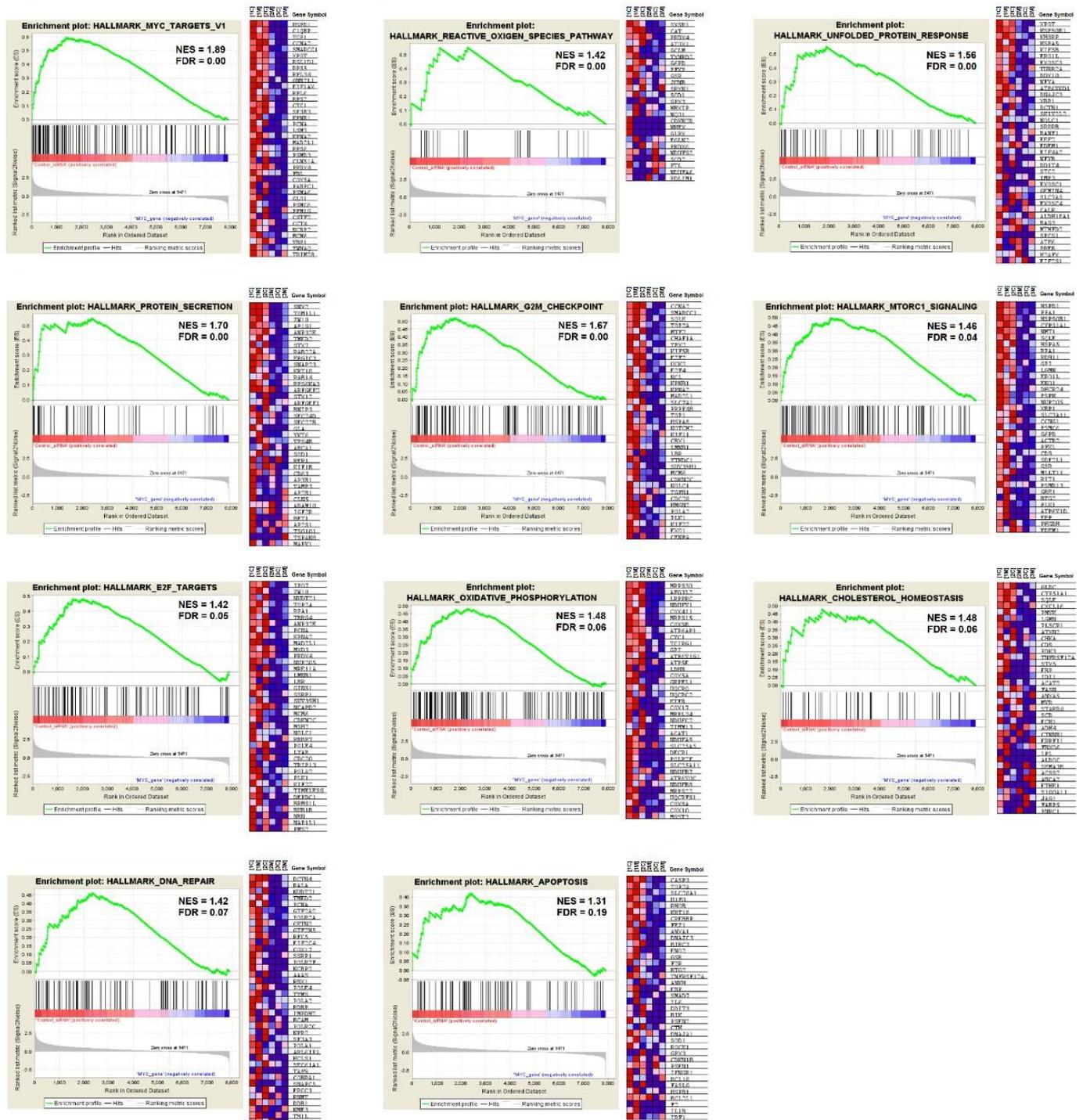


Figura 6. Painel com 11 conjuntos de genes enriquecidos para três linhagens de células de CG após o silenciamento do *MYC*. Cada conjunto foi representado no Enplot pela sua Pontuação de Enriquecimento Normalizado (NES) e Taxa de Descoberta Falsa (FDR), juntamente com o seu mapa de calor mostrando a expressão gênica para cada gene.

5.2 Linhagens Celulares de CG Silenciadas por *MYC* de Diferentes Subtipos Histológicos Mostram Perfis Distintos de Expressão Dependentes de *MYC*

Posteriormente, foram avaliadas as vias metabólicas mais prováveis de serem reguladas por *MYC* como demonstrado por GSEA, listando esses genes nas categorias descritas como padrões de expressão gênica (LIBERZON et al., 2015) e foi observada relação com a proliferação celular, via, sinalização celular, processos metabólicos e resposta a danos no DNA (Tabela 3). Os conjuntos de dados brutos obtidos das linhagens celulares de CG foram refinados e os DEGs foram identificados para cada um dos 11 conjuntos de genes para a coleção H que foram enriquecidos para o fenótipo controle de siRNA. A figura 7 apresenta as classificações para DEGs enriquecidos quando comparados à coleção H com seu escore de enriquecimento médio (ES) para cada análise. Todos os genes enriquecidos da coleção H estão no ANEXO A.

A linhagem celular do tipo difuso (ACP02) apresentou 149 DEGs enriquecidos por GSEA, três deles com expressão aumentada e 146 com expressão reduzida. Genes enriquecidos com níveis acima de 15.000 se ajustaram no perfil de expressão diminuída e foram enfatizadas por conjuntos de genes como alvo *MYC* V1, secreção proteica e via de Espécies Reativas de Oxigênio (ROS) (Figura 7A); por outro lado, níveis acima de 30.000 foram enriquecidos para DEGs com expressão aumentada, apresentando conjuntos de genes descritos como secreção proteica, alvo *MYC* V1, resposta a proteínas mal enoveladas e via de ROS (Figura 7B). Enquanto as células do tipo intestinal (ACP03) demonstraram 85 DEGs, 76 deles com expressão aumentada e nove com expressão diminuída. Genes com níveis acima de 20.000 se ajustaram no perfil de expressão aumentada por vias como secreção proteica, alvo *MYC* V1 e vias de ROS (Figura 7C); genes com expressão diminuída que tiveram níveis acima de 15.000 apresentaram ênfase de enriquecimento para GM2-Checkpoint, secreção proteica, alvo *MYC* V1, via de ROS e resposta a proteínas mal enoveladas (Figura 7D). Para as amostras metastáticas (AGP01), 65 DEGs enriquecidos foram identificados por GSEA, 25 com expressão aumentada e 40 com expressão diminuída, e os conjuntos de genes foram enriquecidos com classificações acima de 20.000 para ambos os perfis DEGs. Os conjuntos que apresentaram expressão aumentada (Figura 7E) foram enriquecidos para GM2-Checkpoint, secreção proteica, alvo *MYC* V1 e via de ROS, enquanto os perfis com expressão diminuída tiveram ênfase na fosforilação oxidativa, alvo *MYC* V1, secreção proteica e via de ROS (Figura 7F).

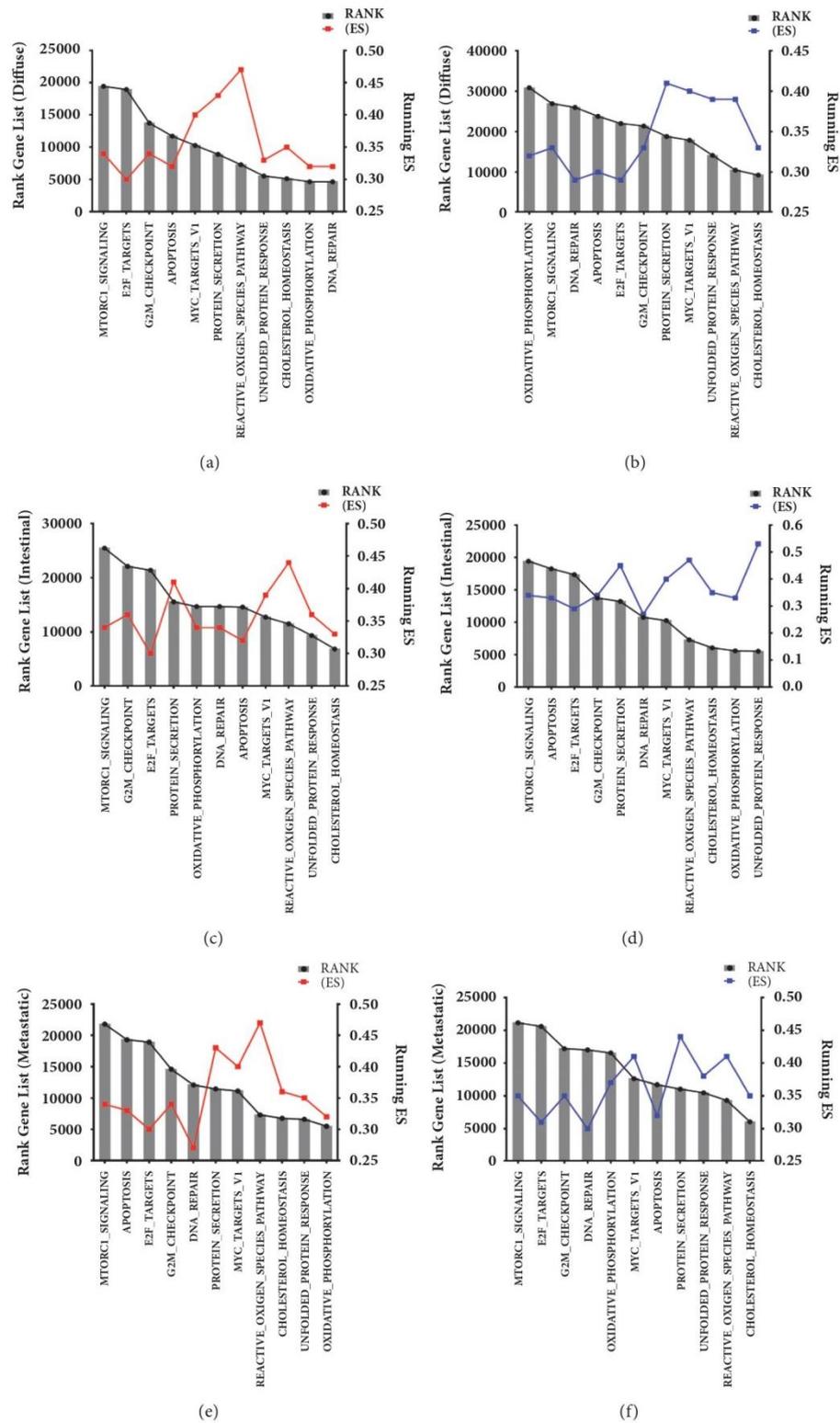


Figure 7. Escores de DEGs enriquecidos quando comparados com a coleção H com seu escore médio de enriquecimento (ES) para cada análise. **(A)** DEGs de ACP02 com expressão aumentada, classificações (4.640 a 19.433) e ES (0,30 a 0,47). **(B)** DEGs de ACP02 com expressão diminuída, classificações (9,245 a 30,875) e ES (0,29 a 0,41). **(C)** DEGs de ACP03 com expressão aumentada, posições (6,853 a 25,475) e ES (0,30 a 0,47). **(D)** DEGs de ACP03 com expressão diminuída, ranks (5,530 a 19,433) e ES (0,27 a 0,53). **(E)** DEGs de AGP01 com expressão aumentada, classificações (5.480 a 21.804) e ES (0,27 a 0,47). **(F)** DEGs de AGP01 com expressão diminuída, ranks (6.060 a 21.134) e ES (0,30 a 0,44).

Nos 11 conjuntos de genes da coleção H usados para o enriquecimento de DEGs, verificaram-se genes com maior ES. Essa estratégia foi utilizada para identificar genes semelhantes que se sobrepõem como um agrupamento de consenso, como descrito anteriormente (LIBERZON et al., 2015), por isso seria mais fácil identificar genes regulados por *MYC*. DEGs identificados para as três linhagens de células de CG que foram enriquecidas quando comparadas à coleção H, com níveis acima de 2.000. O gene alvo *MYC* V1 foi identificado como o mais enriquecido para as linhagens celulares estudadas com maiores escores no ES para os genes *SNRPD2* e *TYMS*. Quando comparamos os genes enriquecidos encontrados para a coleção H para ACP02 e ACP03, notamos que alguns genes apresentaram graus de enriquecimento que também eram comuns para as células metastáticas (AGP01) (Figura 8A-C).

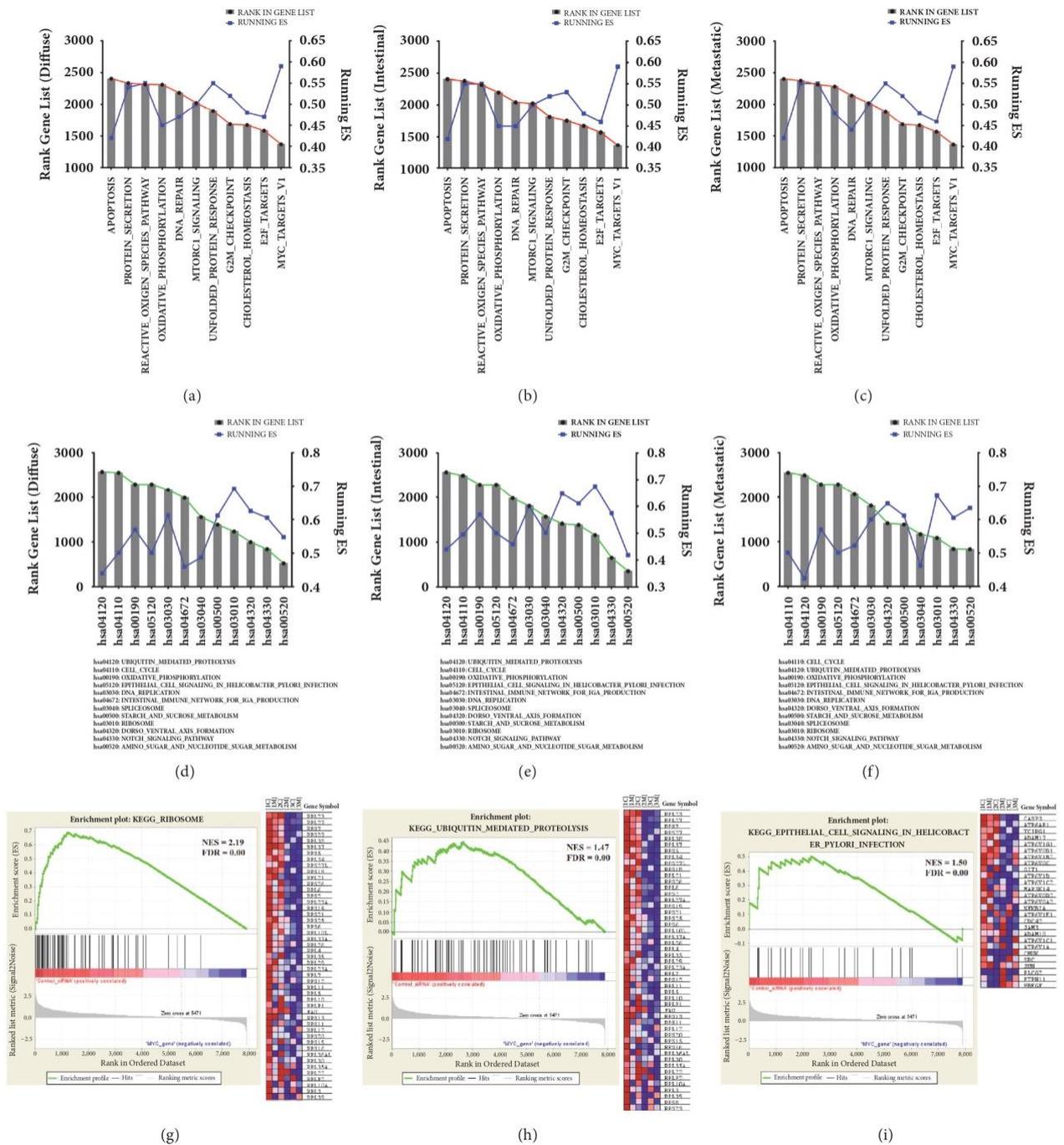


Figure 8. Identificação de genes individuais nos conjuntos de dados DEGs enriquecidos com o ES mais alto. (A) enriquecimento de DEGs para ACP02. (B) enriquecimento de DEGs para ACP03. (C) enriquecimento de DEGs para AGP01. A marca mais enriquecida para as três linhagens celulares como alvo *MYC* V1 com um ES = 0,59. (D) KEGG Identificação da via para ACP02. (E) KEGG identificação da via para ACP03. (F) Identificação KEGG via para AGP01. Enplot e mapa de calor para as vias com as maiores pontuações de ES entre as três linhagens de células de CG podem ser vistas em (G) para ACP02, em (H) para ACP03 e em (I) para AGP01.

A mesma análise foi aplicada para identificar DEGs enriquecidos encontrados para as vias KEGG para aprofundar a compreensão sobre os genes envolvidos na carcinogênese do estômago relacionados ao *MYC*. Os mapas de enriquecimento para KEGG (KANEHISA et al., 2012) são mostrados na Figura 8D-F, apresentando 12 vias significativamente enriquecidas encontradas utilizando a ferramenta GSEA, sugerindo que o CG possui múltiplas vias alteradas que levam a mucosa gástrica normal ao processo carcinogênico. Essas vias foram definidas por um Escore de Enriquecimento Normalizado (NES) > 1,47.

A estatística t média dos genes foi calculada para cada via KEGG usando testes de permutação com 100 repetições. O gráfico de enriquecimento para três vias comuns alteradas em todas as nossas linhagens celulares é apresentado nas seguintes figuras: Figura 8G (ribossomo - hsa03010), que apresentou o maior ES entre as células, Figura 8H (proteólise mediada por ubiquitina - hsa04120), com as mais altas classificações em nossa lista de genes e a Figura 8I (sinalização celular epitelial na infecção por *H. pylori* - hsa05120). A Tabela 2 (ANEXO) apresenta mais detalhes sobre os DEGs enriquecidos com KEGG.

Foram identificados DEGs enriquecidos comuns às três linhagens celulares de CG utilizadas neste estudo e estão representados como diagramas de Venn construídos pela plataforma InteractiVenn (HEBERLE, 2015). Foram encontrados 14 DEGs que são compartilhados enriquecidos nas linhagens celulares AGP01, ACP02 e ACP03 (Figura 9A), que provavelmente são alvos de *MYC* de acordo com a nossa análise. A função da maioria dos genes (14,78%) em diferentes bancos de dados (Figura 9B) está envolvida na secreção proteica, seguida pela resposta a proteínas mal enoveladas (14,39%) e via de ROS (13,91%). A Tabela 3 (ANEXO) apresenta detalhes sobre esses resultados.

A Tabela 4 mostra os 14 DEGs individuais, bem como as características de expressão gênica em que estão envolvidos, classificações e ES. Mostramos um mapa de calor (Figura 9C) dos 14 DEGs compartilhados em clusters agrupados usando a ferramenta GENE E (<https://software.broadinstitute.org/GENE-E/>), onde a expressão diminuída é expressa em azul e a expressão aumentada em vermelho.

Tabela 4. Os 14 DEGs comuns para as três linhagens de células de CG após o MYC-siRNA e as características dos genes que foram enriquecidos em CG.

Nome da característica:	Gene	Classificação na lista de Genes	Pontuação Métrica	Corrida (ES)
SECREÇÃO PROTEICA	<i>BNIP3</i>	1438	6.0E-01	0.522
RESPOSTA A PROTEÍNAS MAL ENOVELADAS	<i>SKIV2L2</i>	1051	7.2E-01	0.508
RESPOSTA A PROTEÍNAS MAL ENOVELADAS	<i>SRPRB</i>	1442	6.0E-01	0.506
VIA DE ESPÉCIES REATIVAS DE OXIGÊNIO	<i>JUNB</i>	1969	4.8E-01	0.491
G2M CHECKPOINT	<i>LBR</i>	817	8.2E-01	0.451
FOSFORILAÇÃO OXIDATIVA	<i>NDUFV2</i>	1542	5.7E-01	0.444
APOPTOSE	<i>CDKN1B</i>	2405	4.0E-01	0.420
SECREÇÃO PROTEICA	<i>RAB22A</i>	289	1.2E+15	0.408
HOMEOSTASE DE COLESTEROL	<i>ATXN2</i>	945	7.6E-01	0.382
HOMEOSTASE DE COLESTEROL	<i>LGMN</i>	445	1.1E+16	0.367
FOSFORILAÇÃO OXIDATIVA	<i>ACAT1</i>	840	8.1E-01	0.367
SECREÇÃO PROTEICA	<i>TMED2</i>	254	1.3E+16	0.326
REPARO DE DNA	<i>AAAS</i>	1109	7.0E-01	0.313
G2M CHECKPOINT	<i>NCL</i>	328	1.2E+13	0.289

ES: escore de enriquecimento.

A maioria dos genes para ACP02 (tipo difuso CG) apresentou expressão diminuída, enquanto para ACP03 (tipo intestinal CG) os mesmos genes apresentaram expressão aumentada. A linhagem celular metastática AGP01, apesar de seu tumor original ser do tipo intestinal, apresentou padrões de expressão mista com expressão diminuída predominante.

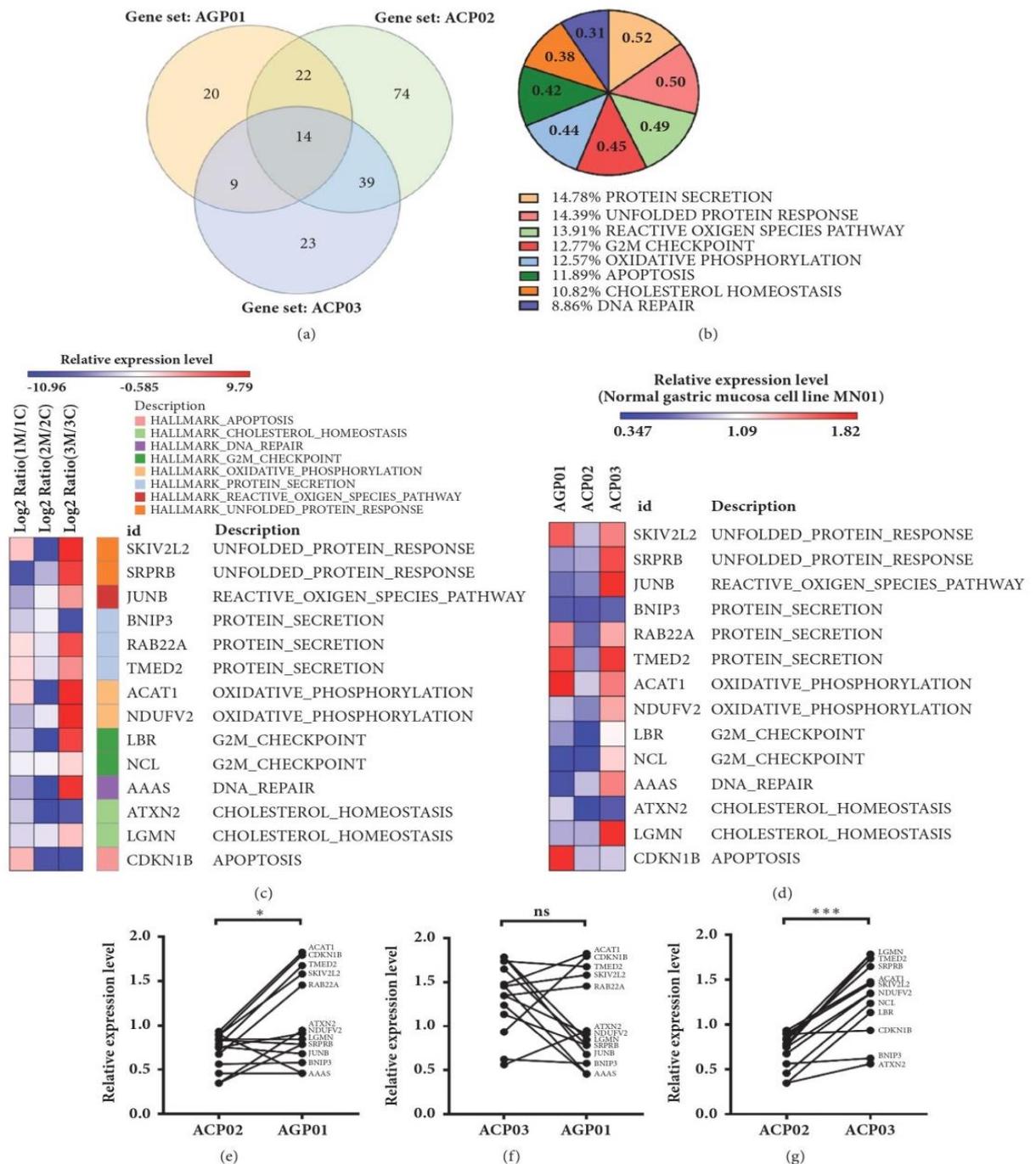


Figura 9. DEGs comuns enriquecidos em três linhagens celulares de CG após o silenciamento do *MYC*. (A) O diagrama de Venn identificou os 14 DEGs comuns entre AGP01, ACP02 e ACP03, mas cada um deles também possui um conjunto exclusivo de DEGs. (B) Identificação dos 14 DEGs comuns em 8 marcas enriquecidas da coleção H, representada pela % de seus correspondentes ES, ES pontuações são mostradas dentro do gráfico de pizza para cada marca. (C) Mapa de calor para a expressão dos 14 DEGs comuns compartilhados pelas três linhagens de CG após o silenciamento do *MYC*. Azul representa a expressão diminuída enquanto o vermelho significa a expressão aumentada. Observe que, mesmo para os 14 genes comuns, cada linhagem celular tem um padrão de expressão diferente. A expressão de leitura relativa foi normalizada usando Log₂ FC entre *MYR* silenciado / controle siRNA. (d) Mapa de calor para a expressão gênica relativa por RT-qPCR para os 14 DEGs identificados para as linhagens AGP01, ACP02 e ACP03. (e) Expressão global para os 14 DEGs é diminuída em ACP02 quando comparada à AGP01. (f) A linhagem ACP03 apresentou níveis de expressão gênicas aumentadas, mas não foram encontradas diferenças estatísticas. (g) Expressão gênica relativa aumentada em ACP03 quando comparada com ACP02. O teste de Wilcoxon foi usado para comparar os níveis relativos de expressão gênica (*P < 0.05, ns: não significante e ***P < 0.0001).

5.3 Os 14 DEGs Regulados por MYC Mostram Perfis de Expressão Distintos para cada Subtipo Histológico de CG

A análise *in silico* identificou um conjunto central de 14 DEGs (Figura 9A) que são regulados por *MYC*, porém o perfil de expressão é distinto para cada linhagem, pois representam diferentes subtipos histológicos. Validamos os perfis de expressão gênica apresentados por AGP01, ACP02 e ACP03 após silenciamento do *MYC* por RT-qPCR no cDNA obtido originalmente para cada linhagem celular e os comparamos com sua expressão na linhagem celular MN01. Os resultados apontam que o conjunto de DEGs identificados está sob a regulação transcricional do *MYC* (Figura 9D). Os resultados da linhagem AGP01 foram heterogêneas, com alguns genes apresentando expressão diminuída e outros apresentando expressão aumentada. A linhagem celular ACP02 apresentou principalmente expressão diminuída na expressão dos 14 DEGs (Figura 9C); por outro lado, ACP03, os mesmos 14 genes, apresentaram-se supra-regulados; nossos resultados de expressão gênica (Figura 9D) confirmam nossa análise *in silico* (Figura 9C).

A comparação da expressão de mRNA foi medida por RT-qPCR para os DEGs identificados afim de avaliar se a expressão gênica desses 14 genes era suficiente para distinguir estatisticamente cada linhagem celular. Ao comparar o ACP02 com o AGP01 (Figura 9E), notamos uma expressão diminuída significativa da expressão gênica para ACP02; confrontando ACP03 versus AGP01 indicou um aumento na quantificação relativa de mRNA (Figura 9F) para ACP03; no entanto, esses resultados não foram significativos; a comparação entre ACP02 versus ACP03 confirmou que os 14 DEGs regulados por *MYC* são significativamente mais expressos em ACP03, a linhagem celular do tipo intestinal, do que em ACP02 (Figura 9G).

6 DISCUSSÃO

Um dos principais objetivos dos estudos sobre o câncer é caracterizar sistematicamente os mecanismos celulares e moleculares envolvidos na doença e seus estágios distintos, para identificar tanto os potenciais biomarcadores quanto os novos prováveis alvos de drogas (HEBERLE et al., 2015).

O perfil molecular do câncer gástrico é heterogêneo, em parte devido a diferentes sistemas de classificação (CHO, 2013) e, para esclarecer as verdadeiras origens moleculares da CG, tanto o The Cancer Genome Atlas (KARIMI et al., 2014) quanto o Asian Cancer Research Group (RUGGE, FASSAN e GRAHAM, 2015) publicaram os subtipos moleculares de câncer gástrico, com notável sobreposição entre os dois modelos. Portanto, vários genes têm sido implicados como biomarcadores para subtipos de CG, como *RHOA*, *EGFR*, *PDL*, *CDH1*, *TP53* e *JAK2*. No entanto, esses estudos utilizam amostras de populações em que a incidência da doença é maior e poucos estudos examinaram populações em que a incidência dessa doença é menor, como no caso do Brasil (CRISTESCU et al., 2015). Há evidências de que a incidência de CG varia entre os países, em grande parte devido à heterogeneidade genética exibida por populações humanas (BINATO et al., 2018), e já foi demonstrado que existe uma assinatura exclusiva de expressão gênica para casos brasileiros de CG do tipo intestinal (CRISTESCU et al., 2015). O presente estudo ajuda a destacar os perfis moleculares das linhagens brasileiras de CG, o que pode ajudar muito nosso entendimento sobre as bases moleculares do CG na América do Sul.

A maioria dos estudos NGS investiga o CG comparando tumor versus tecido não tumoral, analisando padrões globais de expressão gênica, variação no número de cópias e outras características moleculares, entretanto grande parte dos estudos de alto rendimento acerca da genética do câncer gástrico não dão a devida importância ao *MYC* (WALZ e EILERS, 2013; MCLEAN e EL-OMAR, 2014; WANG et al., 2014; KARIMI et al., 2014; BASS et al., 2014; FIGUEIREDO et al., 2017). Os resultados deste estudo são relevantes porque a superexpressão de *MYC* é uma descoberta chave em amostras brasileiras de CG (BUFFART et al., 2011). A expressão desse gene foi reduzida usando siRNA para identificar a assinatura relacionada a *MYC* em linhagens de células de CG, comparando amostras silenciadas versus não silenciadas. Um total de 5.471 DEGs foi identificado e 11 conjuntos significativos de genes, incluindo alvos clássicos de *MYC* representados pelo alvo *MYC V1*.

A análise computacional de conjuntos de genes identificados deu-se após o

silenciamento do *MYC* em linhagens brasileiras de CG (CALCAGNO et al., 2010; DA COSTA et al., 2011), que carregam alterações genéticas comumente encontradas em pacientes brasileiros com CG (LEAL et al., 2009; CALCAGNO et al., 2010; CALCAGNO et al., 2013). Este oncogene promove o crescimento celular atuando como um fator de transcrição que regula o ciclo celular, o metabolismo e a sobrevivência celular (LEAL et al., 2016). Verificou-se o aumento da expressão de DEGs apenas para a linhagem celular do tipo intestinal (ACP03), enquanto o tipo difuso (ACP02) e as células metastáticas de CG (AGP01) apresentaram expressão global da expressão gênica diminuída; ao olhar para genes individuais entre ACP02 e AGP01, ainda é possível distingui-los pela expressão gênica relacionada ao *MYC*. É importante notar que o *MYC* tem um papel duplo no processo carcinogênico, ativando seletivamente e inativando diferentes conjuntos de genes (WALZ et al., 2014; KRESS, SABÒ e AMATI, 2015; POSTERNAK e COLE, 2016). Em conjunto, apresentam-se evidências que confirmam o fato de que a desregulação do *MYC* tem um papel importante na carcinogênese gástrica (LEAL et al., 2011) e que as assinaturas relacionadas ao *MYC* no câncer gástrico são diferentes para cada subtipo histológico desta doença o que é clinicamente relevante (CRISTESCU et al., 2015; KRESS et al., 2016).

Uma das principais formas de regulação da proteína *MYC* em células normais é através de sua degradação direcionada pelo sistema ubiquitina-proteassoma (ZANG et al., 2014), que foi uma das vias enriquecidas em KEGG encontradas nesta análise (Figura 8). Isso significa que, em uma condição de superexpressão de *MYC*, como no CG (CALCAGNO et al., 2013), não apenas este gene e sua proteína são mais produzidos, eles também são menos destruídos porque diminui a expressão de ubiquitinas-ligases E3, como Fbw7 e HectH9, contribuindo para o prolongamento da meia-vida da proteína *MYC* e a amplificação de seus efeitos (FARRELL e SEARS, 2014). As ubiquitinas-ligases, incluindo o Fbw7 reguladas por *MYC*, foram identificadas recentemente como alvos terapêuticos promissores para o desenvolvimento de novos fármacos antineoplásicos (KING et al., 2013).

Outras vias envolvidas na carcinogênese gástrica relacionada ao *MYC* encontradas neste estudo são alvos conhecidos, como o ribossomo e genes de controle do ciclo celular, que dificilmente são alvos terapêuticos. Quando consideradas em conjunto, as 12 diferentes vias encontradas sob o controle de *MYC* para a carcinogênese gástrica representam muitas funções biológicas, o que significa que a superexpressão do *MYC* no CG perturba quase todos os processos celulares regulares em favor do desenvolvimento do tumor (KRESS, SABÒ e AMATI, 2015).

Outro conjunto interessante de genes inclui o metabolismo da glicose, que é uma área

de crescente interesse na pesquisa do câncer, e tem sido demonstrado que *MYC* direciona a ativação da glicólise aeróbica, uma característica do metabolismo do câncer conhecido como efeito Warburg, e praticamente todos os genes envolvidos na glicólise e a maioria dos responsáveis pela glutaminólise (DANG, 2013; LIU et al., 2015).

Também foram identificados 14 DEGs enriquecidos em todas as três linhagens celulares de CG usadas neste estudo que podem representar o conjunto comum de genes regulados por *MYC* na carcinogênese gástrica (Tabela 4). Isso é importante porque mostra que, embora se tenha caracterizado os diferentes subtipos histológicos e os estágios de CG, ainda pouco se sabe sobre este conjunto de genes regulados pelo *MYC* envolvidos no processo carcinogênico. Também não encontramos outros relatos na literatura sobre esses 14 genes e análise de alto rendimento do câncer gástrico (CHO, 2013; BASS et al., 2014; ZANG et al., 2014; WONG et al., 2014; COOLS-LARTIGUE, BAKER e FERRI, 2015; CRISTESCU et al., 2015; KANDA e KODERA, 2015; LIN, 2015; RUGGE, FASSAN e GRAHAM, 2015; BRIA et al., 2016; GARATTINI et al., 2017; KATONA e RUSTGI, 2017; SUN et al., 2017). Mesmo quando se consideram as assinaturas moleculares apresentadas pelo CG do tipo intestinal brasileiro (BINATO et al., 2018) não foram encontradas nenhuma concordância para os 14 genes encontrados neste estudo, mas é importante considerar que Binato e colaboradores (2018) não levaram em consideração a superexpressão de *MYC* em suas amostras. Portanto, esses DEGs encontrados representam novos e importantes achados referentes ao processo da carcinogênese gástrica regulado por *MYC* na população brasileira.

Embora estudos adicionais sejam necessários para validar esses resultados, apresentam-se fortes evidências de que genes regulados por *MYC* em CG têm diferentes padrões de expressão quando consideramos diferenças histológicas e de estágio de doença, no entanto, eles ainda compartilham vias e genes centrais envolvidos no processo carcinogênico.

7 CONCLUSÃO

Nossos resultados identificaram nas linhagens de câncer gástricos analisadas genes que têm sua expressão alterada com o silenciamento do gene *MYC* mostrando que é possível discernir os subtipos histológicos de câncer gástrico por meio da análise do padrão de expressão gênica regulada pelo gene *MYC*. Considerando que o CG possui múltiplas vias celulares alteradas, é esperado que a forma como a proteína *MYC* causa a tumorigênese do estômago seja diferente para cada subtipo da doença.

Esses dados são importantes para a caracterização da ação do gene *MYC* na tumorigênese do câncer gástrico, podendo servir de base para mais estudos que auxiliem no diagnóstico mais específico e na determinação de novos alvos terapêuticos para esta doença.

8 REFERÊNCIAS

- ABDI, H. **The Bonferroni and Sidák Corrections for multiple comparisons**. pp. 1– 9, 2007.
- ALBERTS, B.; BRAY, D.; HOPKIN, K.; JOHNSON, A.; LEWIS, J.; RAFF, M.; ROBERTS, K.; WALTER, P. Comunidades celulares: tecidos, células-tronco e câncer. *In*: Alberts et al. **Fundamentos da biologia celular**. 4ª ed. Porto Alegre: Editora Artmed. pp. 683-720, 2017.
- AUDIC, S.; CLAVERIE, J. **The significance of digital gene expression profiles**. pp. 986–995, 1997.
- BARBOSA, A. S.; LIN, C. J. Silenciamento de genes com RNA de Interferência: Um Novo Instrumento Para Investigação da Fisiologia e Fisiopatologia do Córtex Adrenal. **Arq Bras Endocrinol Metab** vol 48 nº 5 Outubro 2004.
- BARRETINA, J.; CAPONIGRO, G.; STRANSKY, N.; VENKATESAN, K.; MARGOLIN, A.A. et al The Cancer Cell Line Encyclopedia enables predictive modeling of anticancer drug sensitivity. **Nature** 483: 603-607, 2012.
- BASS A. J. et al.* Comprehensive molecular characterization of gastric adenocarcinoma. **Nature**, vol. 513, no. 7517, pp. 202–209, 2014.
- BEHJATI, S.; TARPEY, P. S. What is next generation sequencing? **Arch. Dis. Child Educ. Pract.** ed;98:236–238. doi:10.1136/archdischild-2013-304340, 2013.
- BENJAMINI, Y.; YEKUTIELI, D. The control of the false discovery rate in multiple testing under dependency. **Ann. Stat.**, vol. 29, no. 4, pp. 1165–1188, 2001.
- BERMÚDEZ, C.; INSUASTY, J.; GAMARRA, G. Grupo sanguíneo A y riesgo de cáncer gástrico en el Hospital Universitario de Santander (Bucaramanga, Colombia). **Acta Médica Colombiana**, vol. 31, nº 4, 2006.
- BINATO, R.; SANTOS, E. C.; BORONI, M.; DEMACHKI, S.; ASSUMPCÃO, P.; ABDELHAY, E. A common molecular signature of intestinal-type gastric carcinoma indicates processes related to gastric carcinogenesis. **Oncotarget**, vol. 9, no. 7, pp. 7359–7371, Jan. 2018.
- BRIA, E.; PILOTTO, S.; SIMBOLO, M.; FASSAN, M.; DE MANZONI, G.; CARBOGNIN, L.; SPERDUTI, I.; BRUNELLI, M.; CATALDO, I.; TOMEZZOLI, A.; MAFFICINI, A.; TURRI, G.; KARACHALIOU, N.; ROSELL, R.; TORTORA, G.; SCARPA, A. Comprehensive molecular portrait using next generation sequencing of resected intestinal-type gastric cancer patients dichotomized according to prognosis. **Sci. Rep.**, vol. 6, p. 22982, 2016.
- BROAD INSTITUTE. **GSEA**: gene set enrichment analysis. Disponível em: <<http://software.broadinstitute.org/gsea/index.jsp>>. Acesso em: 01 de Out. 2018.

BUFFART, T. E.; LOUW, M.; VAN GRIEKEN, N. C.; TIJSSEN, M.; CARVALHO, B.; YLSTRA, B.; GRABSCH, H.; MULDER, C. J.; VAN DE VELDE, C. J.; VAN DER MERWE, S. W.; MEIJER, G. A. Gastric cancers of western european and african patients show different patterns of genomic instability. **BMC Med. Genomics**, vol. 4, no. 1, p. 7, 2011.

CALCAGNO, D. Q. ; LEAL, M. F. ; ASSUMPÇÃO, P. P. ; SMITH, M. A. C. ; BURBANO, R. R. . MYC and gastric adenocarcinoma carcinogenesis. **World Journal of Gastroenterology**. v. 14, p. 5962, 2008.

CALCAGNO, D. Q.; LEAL, M. F.; SEABRA, A. D.; KHAYAT, A. S.; CHEN, E. S.; DEMACHKI, S.; ASSUMPÇÃO, P. P.; FARIA, M. H. G.; RABENHORST, S. H. B.; FERREIRA, M. V. P.; SMITH, M. A. C.; BURBANO, R. R. Interrelationship between chromosome 8 aneuploidy, C-MYC amplification and increased expression in individuals from northern Brazil with gastric adenocarcinoma. **World J Gastroenterol**. 12(38): 6207–6211, 2006.

CALCAGNO, D. Q.; LEAL, M. F.; DEMACHKI, S.; ARAÚJO, M. T.; FREITAS, F. W.; OLIVEIRA E SOUZA, D.; ASSUMPÇÃO, P. P.; ISHAK, G.; SMITH, M. A. C.; BURBANO, R. R. MYC in gastric carcinoma and intestinal metaplasia of young adults. **Cancer Genet. Cytogenet.**, vol. 202, no. 1, pp. 63–6, Oct. 2010.

CALCAGNO, D. Q.; FREITAS, V. M.; LEAL, M. F.; DE SOUZA, C. R.; DEMACHKI, S.; MONTENEGRO, R.; ASSUMPÇÃO, P. P.; KHAYAT, A. S.; SMITH, M. A. C.; DOS SANTOS, A. K.; BURBANO, R. R. MYC, FBXW7 and TP53 copy number variation and expression in gastric cancer. **BMC Gastroenterol.**, vol. 13, no. 1, p. 141, 2013.

CHO, J. Y. Molecular diagnosis for personalized target therapy in gastric cancer. **J. Gastric Cancer**, vol. 13, no. 3, pp. 129–135, 2013.

COOLS-LARTIGUE, J.; BAKER, L.; FERRI, L. E. Molecular Mechanisms in Gastric Carcinogenesis. **Gastric Cancer**, Springer, pp. 35–56, 2015.

CRISTESCU, R.; LEE, J.; NEBOZHYN, M.; KIM, K. M.; TING, J. C.; WONG, S. S.; LIU, J.; YUE, Y. G.; WANG, J.; YU, K.; YE, X. S.; DO, I. G.; LIU, S.; GONG, L.; FU, J.; JIN, J. G.; CHOI, M. G.; SOHN, T. S.; LEE, J. H.; BAE, J. M.; KIM, S. T.; PARK, S. H.; SOHN, I.; JUNG, S. H.; TAN, P.; CHEN, R.; HARDWICK, J.; KANG, W. K.; AYERS, M.; HONGYUE, D.; REINHARD, C.; LOBODA, A.; KIM, S.; AGGARWAL, A. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. **Nat. Med.**, vol. 21, no. 5, pp. 449–456, Apr. 2015.

DA COSTA, J. F. F. B.; LEAL, M. F.; SILVA, T. C.; ANDRADE JUNIOR, E. F.; REZENDE, A. P.; MUNIZ, J. A.; LACRETA JUNIOR, A. C.; ASSUMPÇÃO, P. P.; CALCAGNO, D. Q.; DEMACHKI, S.; RABENHORST, S. H.; SMITH, M. D. E. A.; BURBANO, R. R. Experimental Gastric Carcinogenesis in *Cebus apella* Nonhuman Primates. **PLoS One**, vol. 6, no. 7, p. 13, 2011.

DANG, C. V. MYC, metabolism, cell growth, and tumorigenesis. **Cold Spring Harbor Perspectives in Medicine**, vol. 3, no. 8, 2013.

DE SOUZA, C. R. T.; LEAL, M. F.; CALCAGNO, D. Q.; SOZINHO, E. K. C.; BORGES, B. N.; MONTENEGRO, R. C.; DOS SANTOS, A. K. C. R.; DOS SANTOS, S. E. B.; RIBEIRO, H. F.; ASSUMPÇÃO, P. P.; SMITH, M. A. C.; BURBANO, R. R. *MYC* deregulation in gastric cancer and its clinicopathological implications. **PLoS One**, vol. 8, no. 5, 2013.

FARRELL, S.; SEARS, R. C. *MYC* degradation. **Cold Spring Harb. Perspect. Med.**, vol. 4, no. 3, p. a014365, Mar. 2014.

FERNANDEZ, P. C.; FRANK, S. R.; WANG, L.; SCHROEDER, M.; LIU, S.; GREENE, J.; COCITO, A.; AMATI, B. Genomic targets of the human c-MYC protein. **Genes Dev.** 17: 1115–1129, 2003.

FIGUEIREDO, C.; CAMARGO, M. C.; LEITE, M.; FUENTES-PANANÁ, E. M.; RABKIN, C. S.; MACHADO, J. C. **Pathogenesis of gastric cancer: genetics and molecular classification BT - molecular pathogenesis and signal transduction by *Helicobacter pylori***. N. Tegtmeier and S. Backert, Eds. Cham: Springer International Publishing, pp. 277–304, 2017.

FOSTER, I. Cancer: A cell cycle defect. **Radiography** 14 (2): 144-149, 2008.

FOUAD, Y. A; AANEI, C. Revisiting the hallmarks of cancer. **Am J Cancer Res.** 7(5): 1016–1036, 2017.

FRESHNEY, R.I. **Culture of animal cells**. A manual of basic technique and specialized applications. John Wiley & Sons Hoboken, New Jersey. 2010.

GARATTINI, S. K.; BASILE, D.; CATTANEO, M.; FANOTTO, V.; ONGARO, E.; BONOTTO, M.; NEGRI, F. V.; BERENATO, R.; ERMACORA, P.; CARDELLINO, G. G.; GIOVANNONI, M.; PELLA, N.; SCARTOZZI, M.; ANTONUZZO, L.; SILVESTRIS, N.; FASOLA, G.; APRILE, G. Molecular classifications of gastric cancers: Novel insights and possible future applications. **World J. Gastrointest. Oncol.**, vol. 9, no. 5, p. 194, 2017.

GARBER, M.; GRABHERR, M. G.; GUTTMAN, M.; TRAPNELL, C. Computational methods for transcriptome annotation and quantification using RNA-seq. **Nat. Methods**, vol. 8, no. 6, pp. 469–77, Jun. 2011.

GRANDORI, C.; COWLEY, S. M.; JAMES, L. P.; EISENMAN R, N. The *MYC*/Max/Mad network and the transcriptional control of cell behavior. **Annu. Rev. Cell Dev. Biol.**16:653–699, 2000.

GUIMARÃES, R. M.; MUZI, C. D. Trend of mortality rates for gastric cancer in Brazil and regions in the period of 30 years (1980-2009). **Arquivos de Gastroenterologia.** 49(3):184-188, 2012.

HANAHAN, D.; WEINBERG, R. A. **Hallmarks of cancer**: the next generation. **Cell.** v. 144, p. 646-74, 2011.

HARRISON, R. G. **Observations of the living developing nerve fiber**. **Anat. Rec.** 1:116-128; **Proc. Soc. Exp. Med. N.Y.** 140-143, 1907.

HEBERLE, H.; MEIRELLES, G. V.; Da SILVA, F. R.; TELLES, G. P.; MINGHIM, R.

InteractiVenn: a web-based tool for the analysis of sets through Venn diagrams. **BMC Bioinformatics**, vol. 16, no. 1, p. 169, Dec. 2015.

HILLS, S.; DIFFLEY, J. X. DNA Replication and Oncogene-Induced Replicative Stress. **Curr. Biol.**, vol. 24, no. 10, pp. R435–R444, May, 2014.

IARC. (2018.1) **Global cancer observatory**. Disponível em: <<http://gco.iarc.fr/>>. Acesso em: 13 de set. 2018.

IARC. **Cancers attributable to infections**. Disponível em: <http://gco.iarc.fr/causes/infections/toolsbars?mode=2&sex=0&population=who&country=4&continent=0&cancer=0&key=attr_cases&lock_scale=0&nb_results=10>. Acesso em: 02 de out. 2018.

IBQM (UFRJ). **De Sanger ao NGS e de volta**. Disponível em: <<http://www.bioqmed.ufrj.br/docentes/textos/next-generation-sequencing-estado-da-arte/>>. Acesso em: 03 de ago. 2018.

INCA. **ABC do câncer**: abordagens básicas para o controle do câncer. 3. ed. rev. atual. Rio de Janeiro. 108p. 2017.

INCA. **O que é o câncer?** Disponível em: <http://www1.inca.gov.br/conteudo_view.asp?id=322>. Acesso em: 23 de mai. de 2018.

JEMAL, A.; BRAY, F.; CENTER, M. M.; FERLAY, J.; WARD, E.; FORMAN, D. **Global cancer statistics**. CA: A Cancer Journal for Clinicians 61(2): 69-90, 2011.

KANDA M.; KODERA, Y. Recent advances in the molecular diagnostics of gastric cancer. **World J. Gastroenterol. WJG**, vol. 21, no. 34, p. 9838, 2015.

KANEHISA, M.; GOTO, S.; SATO, Y.; FURUMICHI, M.; TANABE, M. KEGG for integration and interpretation of large-scale molecular data sets. **Nucleic Acids Res.**, vol. 40, no. Database issue, pp. D109-14, Jan. 2012.

KARIMI, P.; ISLAMI, F.; ANANDASABAPATHY, S. FREEDMAN, N. D.; KAMANGAR, F. Gastric Cancer: Descriptive epidemiology, risk factors, screening, and prevention. **Cancer Epidemiol. Biomarkers Prev.**, vol. 23, no. 5, pp. 700–713, 2014.

KATONA, B. W.; RUSTGI, A. K. Gastric cancer genomics: advances and future directions. **Cell. Mol. Gastroenterol. Hepatol.**, vol. 3, no. 2, pp. 211–217, 2017.

KING, B.; TRIMARCHI, T.; REAVIE, L.; XU, L.; MULLENDERS, J.; NTZIACHRISTOS, P.; ARANDA-ORGILLES, B.; PEREZ-GARCIA, A.; SHI, J.; VAKOC, C.; SANDY, P.; SHEN, S. S.; FERRANDO, A.; AIFANTIS, I. The ubiquitin ligase FBXW7 modulates leukemia-initiating cell activity by regulating *MYC* stability. **Cell**, vol. 153, no. 7, pp. 1552–1566, 2013.

KRESS, T. R.; PELLANDA, P.; PELLEGRINET, L.; BIANCHI, V.; NICOLI, P.; DONI, M.; RECORDATI, C.; BIANCHI, S.; ROTTA, L.; CAPRA, T.; RAVÀ, M.; VERRECCHIA, A.; RADAELLI, E.; LITTLEWOOD, T. D.; EVAN, G. I.; AMATI, B. Identification of *MYC*-

dependent transcriptional programs in oncogene-addicted liver tumors. **Cancer Res.**, vol. 76, no. 12, pp. 3463–3472, 2016.

KRESS, T. R.; SABÒ, A.; AMATI, B. *MYC*: connecting selective transcriptional control to global RNA production. **Nat. Rev. Cancer**, vol. 15, no. 10, p. 593, 2015.

KU, J. L.; KIM, K. H.; CHOI, J. S.; KIM, S. H.; SHIN, Y. K.; Establishment and characterization of six human gastric carcinoma cell lines, including one naturally infected with Epstein-Barr virus. **Cellular Oncology** 35:127-136, 2012.

LAURÉN, P. The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. An attempt at a histo-clinical classification. **Acta Pathologica Microbiologica et Immunologica Scandinavica** 64: 31-49, 1965.

LEAL, M. F.; CALCAGNO, D. Q.; DA COSTA, J. F. B.; SILVA, T. C.; KHAYAT, A. S.; CHEN, E. S.; ASSUMPÇÃO, P. P.; SMITH, M. A. C.; BURBANO, R. R. *MYC*, *TP53*, and chromosome 17 copy-number alterations in multiple gastric cancer cell lines and in their parental primary tumors. **J. Biomed. Biotechnol.**, vol. 2011, p. 631268, Jan. 2011.

LEAL, M. F.; NASCIMENTO, J. L. M.; DA SILVA, C. E.; LAMARÃO, M. F. V.; CALCAGNO, D. Q.; KHAYAT, A. S.; ASSUMPÇÃO, P. P.; CABRAL, I. R.; SMITH, M. A. C.; BURBANO, R. R. Establishment and conventional cytogenetic characterization of three gastric cancer cell lines. **Cancer Genet. Cytogenet.**, vol. 195, no. 1, pp. 85–91, Nov. 2009.

LEAL, M. F.; WISNIESKI, F.; GIGEK, C. O.; SANTOS, L. C.; CALCAGNO, D. Q.; BURBANO, R. R.; SMITH, M. C. What gastric cancer proteomic studies show about gastric carcinogenesis?. **Tumor Biology**. August, Volume 37, Issue 8, pp 9991–10010, 2016.

LENG, N.; DAWSON, J. A.; THOMSON, J. A.; RUOTTI, V.; RISSMAN, A. I.; SMITS, B. M.; HAAG, J. D.; GOULD, M. N.; STEWART, R. M.; KENDZIORSKI, C. EBSeq: An empirical Bayes hierarchical model for inference in RNA-seq experiments. **Bioinformatics**, vol. 29, no. 8, pp. 1035–1043, 2013.

LI, H. Exploring single-sample SNP and INDEL calling with whole-genome de novo assembly. **Bioinformatics**, vol. 28, no. 14, pp. 1838–1844, Jul. 2012.

LI, H.; DURBIN, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. **Bioinformatics**. vol. 26, no. 5, pp. 589–595, 2010.

LIBERZON, A.; BIRGER, C.; THORVALDSDÓTTIR, H.; GHANDI, M.; MESIROV, J. P.; TAMAYO, P. The molecular signatures database hallmark gene set collection. **Cell Syst.**, vol. 1, no. 6, pp. 417–425, Dec. 2015.

LIMA, E. M.; RISSINO, J. D.; HARADA, M. L.; ASSUMPÇÃO, P. P.; DEMACHKI, S. Conventional cytogenetic characterization of a new cell line, ACP01, established from a primary human gastric tumor. **Brazilian Journal of Medical and Biological Research** 37:1831-1838, 2004.

LIN, Y.; WU, Z.; GUO, W.; LI, J. Gene mutations in gastric cancer: a review of recent next-

- generation sequencing studies. **Tumor Biol.**, vol. 36, no. 10, pp. 7385–7394, 2015.
- LIU, J.; SHAIK, S.; DAI, X.; WU, Q.; ZHOU, X.; WANG, Z.; WEI, W. Targeting the ubiquitin pathway for cancer treatment. **Biochim. Biophys. Acta (BBA)-Reviews Cancer**, vol. 1855, no. 1, pp. 50–60, 2015.
- MA, J.; SHEN, H.; KAPESA, L.; ZENG, S. Lauren classification and individualized chemotherapy in gastric cancer (Review). DOI: 10.3892/ol.2016.4337. **Oncology Letters** 11: 2959-2964, 2016.
- MCLEAN, M. H.; EL-OMAR, E. M. Genetics of gastric cancer. **Nat. Rev. Gastroenterol. Hepatol.**, vol. 11, no. 11, pp. 664–674, 2014.
- MELLO, A. A.; LEAL, M. F.; REY, J. A.; PINTO, G. R.; LAMARÃO, L. M.; MONTENEGRO, R. C.; ALVES, A. P. N.; ASSUMPÇÃO, P. P.; BORGES, B. N.; SMITH, M. C.; BURBANO, R. Deregulated expression of *SRC*, *LYN* and *CKB* Kinases by DNA methylation and its potential role in gastric cancer invasiveness and metastasis. **PLoS ONE** 10(10): e0140492, 2015.
- NAKASHIMA, J.P.; KOIFMAN, R.J.; KOIFMAN, S. Cancer incidence in the Western Amazon: population-based estimates in Rio Branco, Acre State, Brazil, 2007-2009. **Cadernos de Saúde Pública** 28:2125-2132, 2012.
- NAM, H. J.; IM, S. A.; OH, D. Y.; ELVIN, P.; KIM, H. P.; YOON, Y. K.; MIN, A.; SONG, S. H.; HAN, S. W.; KIM, T. Y.; BANG, Y. J. Antitumor activity of saracatinib (AZD0530), a c-Src/Abl kinase inhibitor, alone or in combination with chemotherapeutic agents in gastric cancer. **Mol Cancer Ther.** Jan;12(1):16-26. doi: 10.1158/1535-7163.MCT-12-0109, 2013.
- NCBI. **Gene:** *MYC* proto-oncogene, Bhlh transcription factor [*Homo sapiens* (human)]. Disponível em: <<https://www.ncbi.nlm.nih.gov/gene/4609>>. Acesso em: 20 de set. 2018.
- ORDITURA, M.; GALIZIA, G.; SFORZA, V.; GAMBARDELLA, V.; FABOZZI, A. Treatment of gastric cancer. **World Journal of Gastroenterology** 20(7): 1635-1649, 2014.
- OSTER, S. K.; HO, C. S.; SOUCIE, E. L.; PENN, L. Z. The *MYC* oncogene: MarvelouslyY Complex. **Adv. Cancer Res.** **84**:81–154, 2002.
- PATRO, R.; MOUNT, S. M.; KINGSFORD, C. Sailfish enables alignment-free isoform quantification from RNA-seq reads using lightweight algorithms. **Nat. Biotechnol.**, no. August 2013, pp. 1–6, 2014.
- PEREIRA, T. C. **Introdução à Técnica de Interferência por RNA – RNAi**. Ed.1. SBG - Sociedade Brasileira de Genética. ISBN: 9788589265171. p.170, 2013.
- POSTERNAK, V.; COLE, M. D. Strategically targeting *MYC* in cancer. **F1000Research**, vol. 5, 2016.
- RIVERA, F.; GALÁN, M.; TABERNERO, J.; CERVANTES, A.; VEGA-VILLEGAS, M.E.; GALLEGO, J.; LAQUENTE, B.; RODRÍGUEZ, E.; CARRATO, A.; ESCUDERO, P.; MASSUTÍ, B.; ALONSO-ORDUÑA, V.; CARDENAL, A.; SÁENZ, A.; GIRALT, J.; YUSTE, A. L.; ANTÓN, A.; ARANDA, E. Phase II trial of preoperative irinotecan-cisplatin

followed by concurrent irinotecan-cisplatin and radiotherapy for resectable locally advanced gastric and esophagogastric junction adenocarcinoma. *International Journal of Radiation Oncology. **Biology Physics** 75(5):1430-1436, 2009.*

ROCHA, C. A. M. **As Pernas do Caranguejo**: cancer crura. 1ª. Ed, Belém: Editor Carlos Rocha (ISBN: 978-85-65715-00-3), 114p., 2013.

ROYCHOWDHURY, S.; CHINNAIYAN, A. M. Translating cancer genomes and transcriptomes for precision oncology. *CA Cancer J Clin*;66:75–88, 2016.

RUGGE, M.; FASSAN, M. GRAHAM, D. Y. Epidemiology of gastric cancer. *Gastric Cancer*, Springer, pp. 23–34, 2015.

SILVA, A. M. Ressecções multiorgânicas no câncer gástrico. *Rev Col Bras Cir.* 44(6): 549-552, 2017.

SNUSTAD, D.; SIMMONS, M. J. Base genética do câncer. *In: Fundamentos de genética.* 6ª ed. Rio de Janeiro: Guanabara Koogan, pp. 574-597, 2013.

SUBRAMANIAN, A.; TAMAYO, P.; MOOTHA, V. K.; MUKHERJEE, S.; EBERT, B. L. **Gene set enrichment analysis**: A knowledge-based approach for interpreting genome-wide. 2005.

SUN, C.; YUAN, Q.; WU, D.; MENG, X.; WANG, B. Identification of core genes and outcome in gastric cancer using bioinformatics analysis. *Oncotarget*, vol. 8, no. 41, p. 70271, 2017.

SUZUKI, T.; SEKIGUCHI, M. **Human Cell Culture**: Gastric Cancer. v. 2, p. 257-291, 2002.

THERMO FISHER (2018.1). **Next-Generation Sequencing (NGS)**. Disponível em: <<https://www.thermofisher.com/br/en/home/applications-techniques.html>>. Acesso em: 28 de set. 2018.

THERMO FISHER. **Ion PI™ Chip Kit v3**. Disponível em: <<https://www.thermofisher.com/order/catalog/product/A26771>>. Acesso em: 02 de Out. 2018.

TRAPNELL, C.; WILLIAMS, B. A.; PERTEA, G.; MORTAZAVI, A.; KWAN, G.; VAN BAREN, M. J.; SALZBERG, S. L.; WOLD, B. J.; PACHTER, L. **Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation**. vol. 28, no. 5, 2010.

UICC. **New global cancer data**: GLOBOCAN 2018. Disponível em: <<https://www.uicc.org/new-global-cancer-data-globocan-2018>>. Acesso em: 29 de set. 2018.

VAN NESS, M.; GREGG, J.; WANG, J.; CHEN, M. Genetics and molecular pathology of gastric malignancy: Development of targeted therapies in the era of personalized medicine. *Journal of Gastrointestinal Oncology* 3(3):243-250, 2012.

VITA, M.; HENRIKSSON, M. A. oncoproteína MYC como alvo terapêutico para o câncer

humano. **Semin Cancer Biol.** 16 : 318-330, 2006.

WALZ, S.; EILERS, M. Unlocking the mysterious mechanisms of *MYC*. **Nat. Med.**, vol. 19, no. 1, pp. 26–7, 2013.

WALZ, S.; LORENZIN, F.; MORTON, J.; WIESE, K. E.; VON EYSS, B.; HEROLD, S.; RYCAK, L.; DUMAY-ODELOT, H.; KARIM, S.; BARTKUHN, M.; ROELS, F.; WÜSTEFELD, T.; FISCHER, M.; TEICHMANN, M.; ZENDER, L.; WEI, C. L.; SANSOM, O.; WOLF, E.; EILERS, M. Activation and repression by oncogenic *MYC* shape tumour-specific gene expression profiles. **Nature**, vol. 511, no. 7510, pp. 483–7, 2014.

WANG, K.; YUEN, S. T.; XU, J.; LEE, S. P.; YAN, H. H.; SHI, S. T.; SIU, H. C.; DENG, S.; CHU, K. M.; LAW, S.; CHAN, K. H.; CHAN, A. S.; TSUI, W. Y.; HO, S. L.; CHAN, A. K.; MAN, J. L.; FOGLIZZO, V.; CHING, Y. P.; CHENG, G. H.; XIE, T.; FERNANDEZ, J.; LI, V. S.; CLEVERS, H.; REJTO, P. A.; MAO, M.; LEUNG, S. Y. Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer. **Nat. Genet.**, vol. 46, no. 6, pp. 573–82, 2014.

WEINBERG, R. A. **A biologia do câncer**. 2008. 1ª edição. Porto Alegre: Artmed. 2008.

WHITFIELD, J. R.; SOUCEK, L. Tumor microenvironment: becoming sick of *MYC*. **Cell. Mol. Life Sci.**, vol. 69, no. 6, pp. 931–934, Oct. 2011.

WISNIESKI, F.; CALCAGNO, D. Q.; LEAL, M. F.; DOS SANTOS, L. C.; GIGEK, C. O.; ELIZABETH SUCHI CHEN, E. S.; THAÍS BRILHANTE PONTES, T. B.; ASSUMPCÃO, P. P.; ASSUMPCÃO, M. B.; DEMACHKI, S.; ROMMEL RODRÍGUEZ BURBANO, R. R.; SMITH, M. A. C. Reference genes for quantitative RT-PCR data in gastric tissues and cell lines. **World J. Gastroenterol.** vol. 19, no. 41, pp. 7121–7128, Nov. 2013.

WONG, S. S.; KIM, K. M.; TING, J. C.; YU, K.; FU, J.; LIU, S.; CRISTESCU, R.; NEBOZHYN, M.; GONG, L.; YUE, Y. G.; WANG, J.; RONGHUA, C.; LOBODA, A.; HARDWICK, J.; LIU, X.; DAI, H.; JIN, J. G.; YE, X. S.; KANG, S. Y.; DO, I. G.; PARK, J. O.; SOHN, T. S.; REINHARD, C.; LEE, J.; KIM, S.; AGGARWAL, A. Genomic landscape and genetic heterogeneity in gastric adenocarcinoma revealed by whole-genome sequencing. **Nat. Commun.**, vol. 5, p. 5477, 2014.

YAMASHITA, K.; SAKURAMOTO, S.; WATANABE, M. Genomic and epigenetic profiles of gastric cancer: potential diagnostic and therapeutic applications. DOI 10.1007/s00595-010-4370-5 2011. **Surgery Today**, 41:24–38. 2011.

YANG, Y.; BAI, Z.; YIN, J.; WU, G.; ZHANG, Z. Role of c-Src activity in the regulation of gastric cancer cell migration. DOI: 10.3892/or.2014.3188. **Oncology reports**, 32: 45-49, 2014.

ZALI, H.; REZAEI-TAVIRANI, M.; AZODI, M. Gastric cancer: prevention, risk factors and treatment. **Gastroenterol Hepatol Bed Bench**;4(4):175-185, 2011.

ZANG, S.; GUO, R.; XING, R.; ZHANG, L.; LI, W.; ZHAO, M.; FANG, J.; HU, F.; KANG, B.; REN, Y.; ZHUANG, Y.; LIU, S.; WANG, R.; LI, X.; YU, Y.; CHENG, J.; LU, Y. **Genomics. Proteomics Bioinformatics**, vol. 12, no. 6, pp. 276–283, 2014.

ZHANG, L. L.; HOU, Y.; ASHKTORAB, H.; GAO, L.; XU Y.; WU, K.; ZHAI, J.; ZHANG, L. The impact of C-*MYC* gene expression on gastric cancer cell. **Mol. Cell. Biochem.** vol. 344, no. 1–2, pp. 125–135, Nov. 2010.

ZILBERSTEIN, B.; JACOB, C. E; CECCONELLO, I. Gastric cancer trends in epidemiology. **Arq.Gastroenterol.** vol.49 no.3 São Paulo July/Sept. 2012.

APÊNDICE A - ARTIGO PUBLICADO

Hindawi
 Canadian Journal of Gastroenterology and Hepatology
 Volume 2018, Article ID 5804376, 14 pages
<https://doi.org/10.1155/2018/5804376>



Research Article

Gastric Cancer Cell Lines Have Different *MYC*-Regulated Expression Patterns but Share a Common Core of Altered Genes

Jersey Heitor da S. Maués ¹, Helem Ferreira Ribeiro ^{1,2}, Giovanni R. Pinto,³
 Luana de Oliveira Lopes,³ Letícia M. Lamarão,⁴ Carla Mariana F. Pessoa,⁵
 Caroline de Fátima Aquino Moreira-Nunes,⁶ Raimundo Miranda de Carvalho,⁷
 Paulo P. Assumpção,⁵ Juan A. Rey,⁸ and Rommel M. Rodríguez Burbano⁷

¹Human Cytogenetics Laboratory, Institute of Biological Sciences, Federal University of Pará, Belém, Brazil

²Center of Biological and Health Sciences, Department of Biomedicine, University of Amazon, Belém, Brazil

³Department of Biomedicine, Federal University of Piauí, Parnaíba, Brazil

⁴Laboratory of Nucleic Acids, State Center of Hematology and Hemotherapy, Belém, Brazil

⁵Oncology Research Nucleus, University Hospital João de Barros Barreto, Federal University of Pará, Belém, Brazil

⁶Laboratory of Pharmacogenetics, Drug Research and Development Center, Federal University of Ceará, Fortaleza, Brazil

⁷Laboratory of Molecular Biology, Ophir Loyola Hospital, Belém, Brazil

⁸Molecular Oncogenetics Laboratory, Research Unit, Hospital Universitario La Paz, Madrid, Spain

Correspondence should be addressed to Jersey Heitor da S. Maués; jerseymaues@gmail.com

Received 27 April 2018; Revised 12 September 2018; Accepted 23 September 2018; Published 16 October 2018

Guest Editor: Kiran L. Sharma

Copyright © 2018 Jersey Heitor da S. Maués et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

MYC is an oncogene responsible for excessive cell growth in cancer, enabling transcriptional activation of genes involved in cell cycle regulation, metabolism, and apoptosis, and is usually overexpressed in gastric cancer (GC). By using siRNA and Next-Generation Sequencing (NGS), we identified *MYC*-regulated differentially expressed Genes (DEGs) in three Brazilian gastric cancer cell lines representing the histological subtypes of GC (diffuse, intestinal, and metastasis). The DEGs were picked using *Sailfish* software, followed by Gene Set Enrichment Analysis (GSEA) and Kyoto Encyclopedia of Gene and Genome (KEGG) pathway analysis using KEGG. We found 11 significantly enriched gene sets by using enrichment score (ES), False Discovery Rate (FDR), and nominal P-values. We identified a total of 5,471 DEGs with correlation over (80%). In diffuse-type and in metastatic GC cell lines, *MYC*-silencing caused DEGs downregulation, while the intestinal-type GC cells presented overall DEGs upregulation after *MYC* siRNA depletion. We were able to detect 11 significant gene sets when comparing our samples to the hallmark collection of gene expression, enriched mostly for the following hallmarks: proliferation, pathway, signaling, metabolic, and DNA damage response. When we analyzed our DEGs considering KEGG metabolic pathways, we found 12 common branches covering a wide range of biological functions, and three of them were common to all three cell lines: ubiquitin-mediated proteolysis, ribosomes, and system and epithelial cell signaling in *Helicobacter pylori* infection. The GC cell lines used in this study share 14 *MYC*-regulated genes, but their gene expression profile is different for each histological subtype of GC. Our results present a computational analysis of *MYC*-related signatures in GC, and we present evidence that GC cell lines representing distinct histological subtypes of this disease have different *MYC*-regulated expression profiles but share a common core of altered genes. This is an important step towards the understanding of *MYC*'s role in gastric carcinogenesis and an indication of probable new drug targets in stomach cancer.

1. Introduction

Gastric cancer (GC) remains as an important cause of cancer-related morbidity and mortality worldwide, with recent estimates accounting for over 950,000 new diagnosis and 720,000

deaths each year [1]. Treatment of GC at advanced stages remains difficult, and the prognosis is still poor, partly as a result of local recurrence, tumor invasion, and/or metastasis [2]

The *MYC* oncogene, located at 8q24, is a key oncogene in gastric carcinogenesis, and an increase in both copy number

TABLE 1: Samples used in this study. Control siRNA samples are labeled as C and MYC-siRNA as M.

Cell line	Samples (GEO)	Sample Name	GC Histological subtype
AGP01	GSM2147866	1C	Ascitic fluid of intestinal GC
	GSM2147867	1M	Ascitic fluid of intestinal GC
ACP02	GSM2147868	2C	Diffuse
	GSM2147869	2M	Diffuse
ACP03	GSM2147870	3C	Intestinal
	GSM2147871	3M	Intestinal

and mRNA expression was classified as one of the driver mutations in gastric tumors [3]. *MYC* amplification and overexpression are present in 6-58% of all sporadic gastric tumors [4–6], being more frequent in Brazilian samples [7–9], usually as a result from gene amplification and chromosomal translocations [2, 10].

Our research group previously reported that *MYC* mRNA and protein overexpression is a common finding in GC samples and in some preneoplastic gastric lesions [7, 11–14] from a Brazilian population, as well as in nonhuman primate models of gastric carcinogenesis [15]. We also established and characterized three GC cell lines, AGP01, ACP02, and ACP03, obtained from intestinal-type GC metastasis, diffuse-type GC, and intestinal-type GC, respectively (Leal et al. 2009). Those cell lines also carry genetic alterations commonly found in Brazilian GC patients, such as *MYC* amplification and overexpression and *TP53* deletion [7, 13, 16].

Some consequences of excessive intracellular *MYC* levels are genomic instability [17] and error-prone DNA replication caused by oncogene-induced replicative stress [18]. Even though there is an association between an increase in *MYC* expression and gastric cancer, its exact role in gastric tumorigenesis is not yet fully understood [19, 20] and most of the high-throughput studies carried so far concerning gastric cancer genetics overlook *MYC*'s importance in this process [2, 3, 21–24].

Bioinformatics has mostly been applied in basic science research. Following the completion of human genome sequencing, it has also facilitated numerous discoveries in basic medicine, and several clinical applications of bioinformatics have been reported, including clinical sequencing, an emerging field of precision medicine [25]. In cancer research, bioinformatics has been used to study cancer transcriptome, early diagnosis, cancer grading, and prognosis prediction [26].

In this study, we used RNA interference (RNAi) to block *MYC*'s mRNA translation, followed by Ion Proton™ semiconductor sequencing, in order to identify *MYC*'s regulation signature in AGP01, ACP02, and ACP03 cell lines. We found 11 common pathways for the GC cell lines, which we believe can help in the understanding of expression signatures in different GC histological subtypes.

2. Materials and Methods

2.1. Cell Lines and siRNA Transfection. Three GC cell lines previously established and characterized by our group were used: AGP01, ACP02, and ACP03 [27]. The three cell lines

present chromosome 8 trisomy, *MYC* amplification [13, 27], and *TP53* deletion, which are common genetic alterations in Brazilian gastric cancer patients [28] and in another GC cell line developed in Brazil [29]. A cell culture of nonneoplastic gastric mucosa cells (MNP01, Normal Gastric Mucosa Cell Line 01) pooled from 10 patients without gastric cancer or any other gastric disease, was also used to evaluate the gene and protein expression after *MYC*-silencing and to validate the knockdown results, as well as the *MYC*-regulated genes identified after NGS.

A total of 3×10^5 cells were seeded into 6 cm² plates for each cell line for 24 h before transfection. Small interfering RNAs (siRNA) targeting *MYC* (ON-TARGETplus Human *MYC* (4609) siRNA Dharmacon, EUA) or scrambled control siRNAs (ON-TARGETplus Non-Targeting Pool, Dharmacon, EUA) were transfected into AGP01, ACP02, and ACP03 cell lines using Lipofectamine RNAiMAX Transfection Reagent (Thermo Fisher Scientific, EUA). Optimal transfection was reached after 48 h, and total RNA and proteins were extracted with TRIzol reagent (Thermo Fisher Scientific, EUA). All siRNA experiments were performed three times. The sample names and GEO access codes are shown in Table 1. All siRNA experiments were carried out in biological triplicates.

2.2. Semiconductor Sequencing and Data Pretreatment. Total RNA samples were first treated with DNase-I to remove any possible DNA contamination, and then the mRNA was enriched using Dynabeads Oligo(dt)₂₅ (Thermo Fisher Scientific, USA). The enriched mRNA was fragmented in smaller fragments of 200 bps approximately, which were attached to adapters with known sequences that were unique for each sample. Samples were connected to magnetic beads containing complementary sequences for the adapters and then inserted in microwells where an emulsion-PCR for cDNA synthesis was carried (illustra Ready-To-Go RT-PCR Beads, GE Lifesciences). Our six cDNA libraries were submitted to quantification and quality control using Agilent 2100 Bioanalyzer and were then loaded in Ion Proton V2 PI chip using the Ion PI™ 200 Sequencing Kit v3 and sequenced using Ion Proton™ (Thermo Fisher Scientific, EUA) platform in a single multiplex run.

Raw data reads obtained by primary sequencing using Ion Proton™ were submitted to quality control to calculate alignment and to assess how the reads behave when compared to the reference human genome (Hg19/GRCh37). The aligned reads were mapped and quantified using TMAP (Torrent Mapping Alignment Program), which supports different alignment algorithms [30–32]. Processed datasets were

TABLE 2: Reads quantification after MYC-siRNA.

Sample name	Total Reads	Total Mapped Reads	Gene mapping rate	Expressed genes
1C	13.319.736	12.944.452	97.18%	14.594
1M	11.043.607	10.771.298	97.53%	10.709
2C	12.021.142	11.751.830	97.76%	9.859
2M	11.701.695	11.481.863	98.12%	8.793
3C	12.772.535	12.546.479	98.23%	7.988
3M	12.430.807	12.112.157	97.44%	10.685

uploaded to GEO (Gene Expression Omnibus) under the access number GSE81265.

2.3. Identification and Statistical Analysis of DEGs. Sailfish software pack [33] and the RPKM (Reads Per Kilobase per million mapped reads) [34, 35] were used to significance analysis of DEGs between control and MYC-silenced samples. To identify the DEGs between two paired samples, we used the Audic-Claverie test [36]. Fold-change (FC) was calculated as the \log_2 ratio between the silenced (M) and the control (C) sample. We used a p-value correction corresponding to differential expression tests using Bonferroni correction [37]. Our cut-off for DEGs definition was established as a False Discovery Rate (FDR) < 0.05 [38] and $|\log_2(\text{FC})| > 1$. DEGs were plotted using Multiplot v2, which exhibits personalized gene expression profiles (<http://software.broadinstitute.org/cancer/software/genepattern/modules/docs/multiplot/2>).

2.4. Functional Enrichment Analysis. We used Gene Set Enrichment Analysis (GSEA) [39] to identify significantly enriched gene sets between siRNA control versus MYC-silenced and vice versa. The gene expression changes perceived by DEGs were related to biologically enriched pathways found in GSEA. The gene expression datasets used were collections H (Hallmark gene set) [40] and C2 (curated gene set: KEGG), publicly available at MsigDB [41]. The standard parameters defined by Subramanian et al. were used in our analysis. The statistical significance of GSEA analysis was determined by 100 permutations, the enrichment maps were created to significant ($P < 0.05$ and False Discovery Rate (FDR) < 0.25) gene sets, and GraphPad Prism™ Software was used to graphically represent our data.

2.5. Real-Time Quantitative PCR (RT-qPCR). To confirm the silencing effect of siRNA on MYC expression, we used real-time quantitative PCR to evaluate its expression in relation to the expression found for the normal gastric mucosa cell line MN01. All tests were made in triplicate and using TaqMan® probes as assays-on-demand products for gene expression (Life Technologies, EUA) (MYC: Hs00153408_m1). The fourteen DEGs identified by our analysis (SKIV2L2: Hs00299011_m1, SRPRB: Hs00253639_m1, JUNB: Hs00357891_s1, BNIP3: Hs00969291_m1, RAB22A: Hs00221082_m1, TMED2: Hs00607277_m1, ACATI: Hs00608002_m1, NDUFV2: Hs00221478_m1, LBR: Hs01032700_m1, NCL: Hs01066668_m1, AAAS: Hs00210351_m1, ATXN2:

Hs00268077_m1, LGMN: Hs00271599_m1, and CDKN1B: Hs00153277_m1) were also analyzed and validated by real-time quantitative PCR. The expression of those genes was calculated relative to their expression in the normal gastric mucosa cell line MN01.

3. Results

3.1. Quantification of GCs Transcripts and Identification of DEGs Using NGS. In this study, we used next-generation sequencing based in semiconductors, as well as RNA-Seq, to quantify the transcripts and its isoforms in three gastric cancer cell lines, ACP02, ACP03, and AGP01, before and after MYC-silencing using siRNA. The use of siRNA to reduce MYC expression in the three gastric cancer cell line used in this study was very effective, reducing MYC mRNA expression in 73% for AGP01, in 84% for ACP02, and in 77% for ACP03.

Our NGS sequencing of six libraries generated over 75 million reads, which, after enrichment, were mapped within the reference genome in over 99% of the samples and in over 98% of the reference transcriptome (Table 2), and the distribution of the amplified segments was consistent in all samples. Table 2 also shows in average how many genes were identified for each sample. The average reads produced by Ion Proton™ ranged between 125 and 130 bps.

According to our cut-off (FDR < 0.05 and $|\log_2(\text{FC})| > 1$), we obtained a distinct amount of DEGs between siRNA control versus MYC -silenced samples. Using Multiplot (v2), we identified 1.556 downregulated and 917 upregulated DEGs for AGP01; for the diffuse-type cell line (ACP02), we found 4.098 downregulated versus 1.229 upregulated DEGs; finally, for ACP03, an intestinal-type cell line, we identified 3.272 upregulated versus 842 downregulated DEGs (Figure 1). Our results indicate that it is possible to discern histological subtypes of GC by analyzing its MYC-related gene expression pattern.

A total of 16.777 genes from our six datasets obtained by RNA-Seq (GSE81265) were inserted in an expression matrix normalized by RPKM and, after that, used for enrichment comparing with expression datasets from collections H and C2. By applying GSEA, we looked for gene sets which presented enrichment only in siRNA control samples, but not in MYC -silenced samples, likely MYC targets, and found 11 significant gene sets ($P < 0.05$ and FDR < 0.25), as shown in Table 3. We found a total of 7903 genes, and 5471 (69.2%) are enriched in siRNA control samples, presenting a very high correlation between biological replicates and libraries (80%).

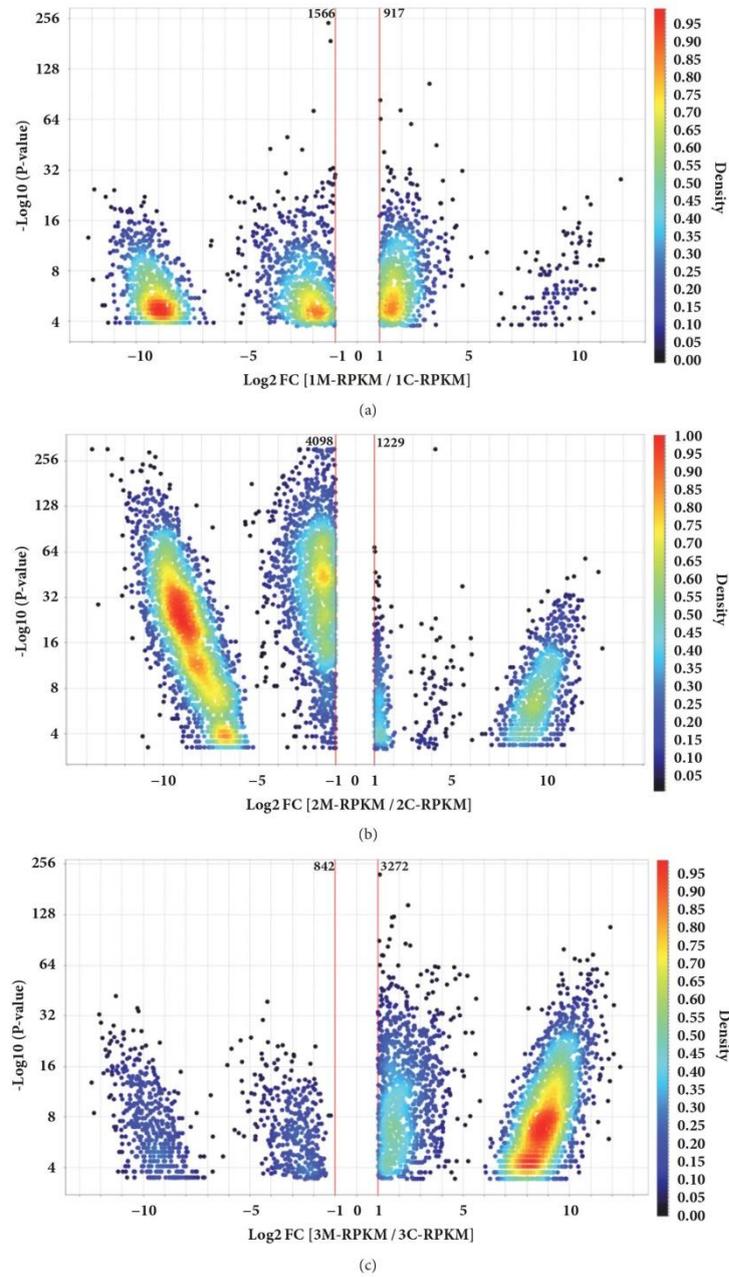


FIGURE 1: Volcano plots of DEGs for three GC cell lines after *MYC*-silencing. \log_2 Fold-change and P-values ($-\log_{10}$) are shown for DEGs with ($|\log_2(\text{FC})| > 1$ and $p \leq 0.05$). (a) DEGs for AGP01. (b) DEGs for ACP02. (c) DEGs for ACP03. Density is a special-case calculation which is only used in advanced shading. It considers only the operands for this calculation are the X and Y values of the graph axes.

TABLE 3: Enriched hallmark gene sets for DEGs after MYC-siRNA in three GC cell lines.

Hallmark Name:	Process category	Description	Size*	ES	NES	Nom. P-value	FDR	Max. Rank
MYC_TARGETS_V1	Proliferation	MYC targets, variant 1	100/44	0.59	1.89	0.00	0.000	1368
REACTIVE_OXIGEN_SPECIES_PATHWAY	Pathway	Reactive oxygen species pathway	25/13	0.55	1.42	0.00	0.000	2316
UNFOLDED_PROTEIN_RESPONSE	Pathway	Unfolded protein response; ER stress	50/22	0.55	1.56	0.00	0.000	1891
PROTEIN_SECRETION	Pathway	Protein secretion	44/26	0.55	1.70	0.00	0.000	2376
G2M_CHECKPOINT	Proliferation	Cell cycle progression: G2/M checkpoint	97/40	0.52	1.67	0.00	0.000	1813
MTORC1_SIGNALING	Signaling	mTORC1 signaling	102/43	0.50	1.46	0.00	0.044	2021
E2F_TARGETS	Proliferation	Cell cycle progression: E2F targets	89/35	0.48	1.42	0.00	0.052	1607
OXIDATIVE_PHOSPHORYLATION	Metabolic	Oxidative phosphorylation and citric acid cycle	118/57	0.48	1.46	0.00	0.062	2284
CHOLESTEROL_HOMEOSTASIS	Metabolic	Cholesterol homeostasis	39/14	0.48	1.48	0.00	0.069	1671
DNA_REPAIR	DNA damage	DNA repair	80/38	0.46	1.42	0.00	0.077	2337
APOPTOSIS	Pathway	Programmed cell death; caspase pathway	66/29	0.42	1.31	0.20	0.190	2405

*Number of genes found for each gene set after curating data. ES: enrichment score. NES: normalized enrichment score. Nom. P-value: nominal P-values for each gene set. FDR: false discovery rate. Max. rank: maximum rank numbers for each gene sets.



FIGURE 2: Panel with 11 gene sets enriched for three GC cell lines after MYC-silencing. Each gene set was represented on the Enplot by its normalized enrichment score (NES) and False Discovery Rate (FDR), along with its heatmap showing the gene expression for each gene.

The enrichment maps obtained are shown in Figure 2, where we created a panel of 11 gene sets significantly enriched.

3.2. MYC-Silenced GC Cells Lines from Different Histological Subtypes Show Distinct MYC-Dependent Expression Profiles. We

then evaluated the metabolic pathways more likely to be MYC-regulated as shown by GSEA, listing those genes in the categories described as hallmarks of gene expression [40] and found that they are related to cellular proliferation, pathway, cellular signaling, metabolic processes, and response to DNA

TABLE 4: The 14 common DEGs for three GC cell lines after *MYC*-siRNA and the gene set hallmarks they were enriched in.

Hallmark Name:	Gene Symbol	Rank in Gene List	Rank Metric Score	Running (ES)
PROTEIN_SECRETION	<i>BNIP3</i>	1438	6.0E-01	0.522
UNFOLDED_PROTEIN_RESPONSE	<i>SKIV2L2</i>	1051	7.2E-01	0.508
UNFOLDED_PROTEIN_RESPONSE	<i>SRPRB</i>	1442	6.0E-01	0.506
REACTIVE_OXYGEN_SPECIES_PATHWAY	<i>JUNB</i>	1969	4.8E-01	0.491
G2M_CHECKPOINT	<i>LBR</i>	817	8.2E-01	0.451
OXIDATIVE_PHOSPHORYLATION	<i>NDUFV2</i>	1542	5.7E-01	0.444
APOPTOSIS	<i>CDKN1B</i>	2405	4.0E-01	0.420
PROTEIN_SECRETION	<i>RAB22A</i>	289	1.2E+15	0.408
CHOLESTEROL_HOMEOSTASIS	<i>ATXN2</i>	945	7.6E-01	0.382
CHOLESTEROL_HOMEOSTASIS	<i>LGMN</i>	445	1.1E+16	0.367
OXIDATIVE_PHOSPHORYLATION	<i>ACAT1</i>	840	8.1E-01	0.367
PROTEIN_SECRETION	<i>TMED2</i>	254	1.3E+16	0.326
DNA_REPAIR	<i>AAAS</i>	1109	7.0E-01	0.313
G2M_CHECKPOINT	<i>NCL</i>	328	1.2E+13	0.289

ES: enrichment score.

damage (Table 3). We refined the raw data sets obtained from our GC cell lines, identifying the DEGs for each of the 11 gene sets for collection H that were enriched for the Control-siRNA phenotype. Figure 3 shows the ranks for enriched DEGs when compared to collection H with their average enrichment score (ES) for each analysis. Part of these results generated in this analysis is in Table 1 (see Supplementary Materials), which show the most enriched DEGs in the set gene *MYC* TARGETS V1.

The diffuse-type cell line (ACP02) presented 149 enriched DEGs by GSEA, 3 up- and 146 downregulated. Genes enriched with ranks above 15,000 fit the downregulated profile and were emphasized for gene sets such as *MYC* target V1, protein secretion, and reactive oxygen species pathway (Figure 3(a)). On the other hand, ranks above 30,000 were enriched for downregulated DEGs, presenting gene sets described as protein secretion, *MYC* target V1, unfolded protein response, and reactive oxygen species pathway (Figure 3(b)). Meanwhile, the intestinal-type cells (ACP03) showed 85 DEGs, 76 up- and 9 downregulated; genes with ranks above 20,000 were upregulated for pathways such as protein secretion, *MYC* target V1, and reactive oxygen species pathway (Figure 3(c)); downregulated genes with ranks above 15,000 presented enrichment emphasis for GM2-checkpoint, protein secretion, *MYC* target V1, reactive oxygen species pathway, and unfolded protein response (Figure 3(d)). For the metastatic samples (AGP01), 65 enriched DEGs were identified by GSEA, 25 up- and 40 downregulated, and the gene sets were enriched with ranks above 20,000 for both DEGs profiles. Upregulated sets (Figure 3(e)) were enriched for GM2-checkpoint, protein secretion, *MYC* target V1, and reactive oxygen species pathway, while the downregulated profile had emphasis for Oxidative.Phosphorylation, *MYC* target V1, protein secretion, and reactive oxygen species pathway (Figure 3(f)).

In the 11 gene sets from collection H that were used for DEGs enrichment, we identified the genes with higher ES. We used this strategy to identify similar genes that overlap as a consensus grouping as described elsewhere [40], so it would

be easier to identify *MYC*-regulated genes. We identified DEGs for the three GC cell lines that were enriched when compared to collection H, with ranks above 2,000. The *MYC* target V1 gene set was identified as the most enriched for our studied cell lines, with higher ES scores for the genes *SNRPD2* and *TYMS*. When we compared the enriched genes found for collection H for ACP02 and ACP03, we noticed that some genes presented enrichment ranks that were also common for the metastatic cells (AGP01) (Figures 4(a)–4(c)).

The same analysis was applied to identify enriched DEGs found for KEGG pathways, to deepen our understanding about the genes involved in the *MYC*-related carcinogenesis for the stomach. The enrichment maps for KEGG [42] are shown in Figures 4(d)–4(f), presenting 12 significantly enriched pathways found using GSEA, suggesting that GC has multiple altered pathways that lead normal gastric mucosa into the carcinogenic process. These pathways were defined by a normalized enrichment score (NES) > 1.47. The average t-statistic for the genes was calculated for each KEGG pathway using permutation tests with 100 repetitions. The enrichment plot for three common pathways altered in all our cell lines is presented in Figure 4(g) (ribosome-hsa03010), which presented the highest ES among the cells, Figure 4(h) (ubiquitin-mediated proteolysis-hsa04120), with the highest ranks in our gene list, and Figure 4(i) (epithelial cell signaling in *Helicobacter pylori* infection-hsa05120). For more details on the DEGs enriched with KEGG, see Table 2 in Supplementary Materials.

We identified enriched DEGs common to all three GC cell lines used in this study and represent them as Venn diagrams constructed by the InteractiVenn platform [43]. We found 14 DEGs that are shared enriched in both AGP01, ACP02, and ACP03 cell lines (Figure 5(a)), which are likely to be *MYC* targets according to our analysis. When we search for gene function in different databases (Figure 5(b)), most of them (14.78%) are involved in protein secretion, followed by unfolded protein response (14.39%) and reactive oxygen species pathway (13.91%). Table 4 shows the 14 individual

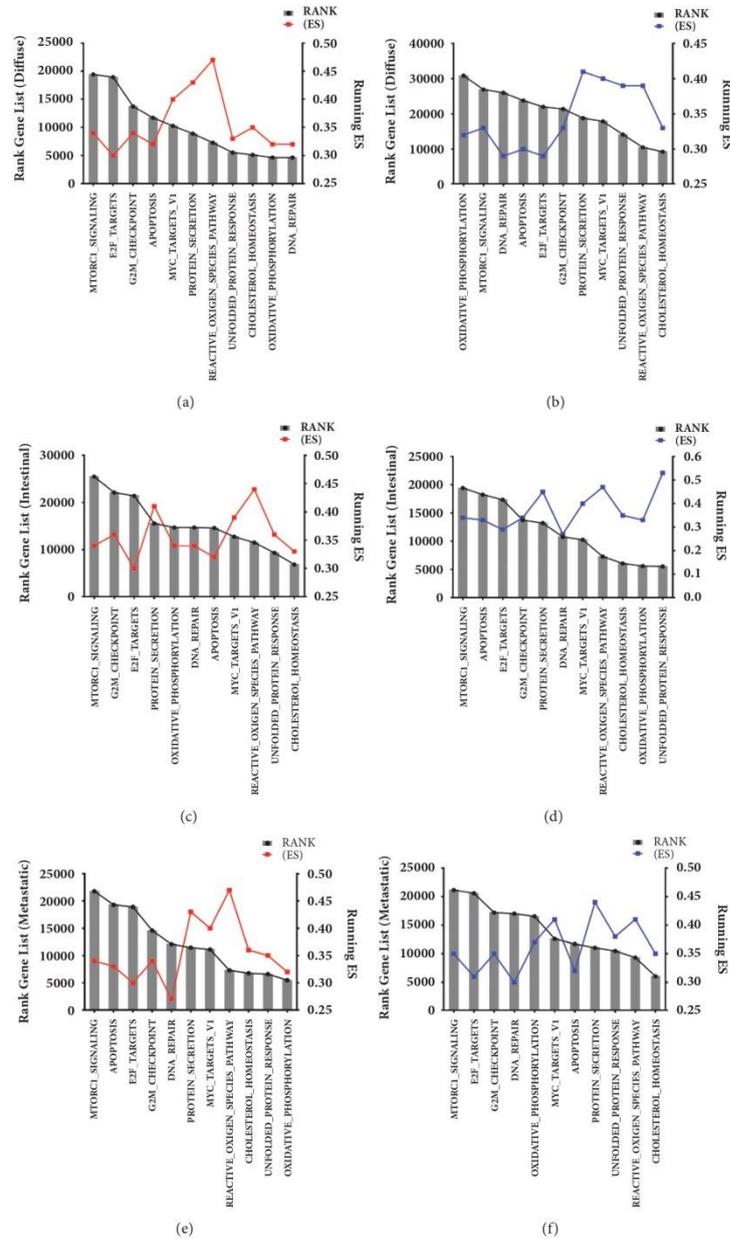


FIGURE 3: Ranks for enriched DEGs when compared to collection H with their average enrichment score (ES) for each analysis. (a) ACP02 upregulated DEGs, ranks (4.640 to 19.433), and ES (0.30 to 0.47). (b) ACP02 downregulated DEGs, ranks (9.245 to 30.875), and ES (0.29 to 0.41). (c) ACP03 upregulated DEGs, ranks (6.853 to 25.475), and ES (0.30 to 0.47). (d) ACP03 downregulated DEGs, ranks (5.530 to 19.433), and ES (0.27 to 0.53). (e) AGP01 upregulated DEGs, ranks (5.480 to 21.804), and ES (0.27 a 0.47). (f) AGP01 downregulated DEGs, ranks (6.060 to 21.134), and ES (0.30 a 0.44).

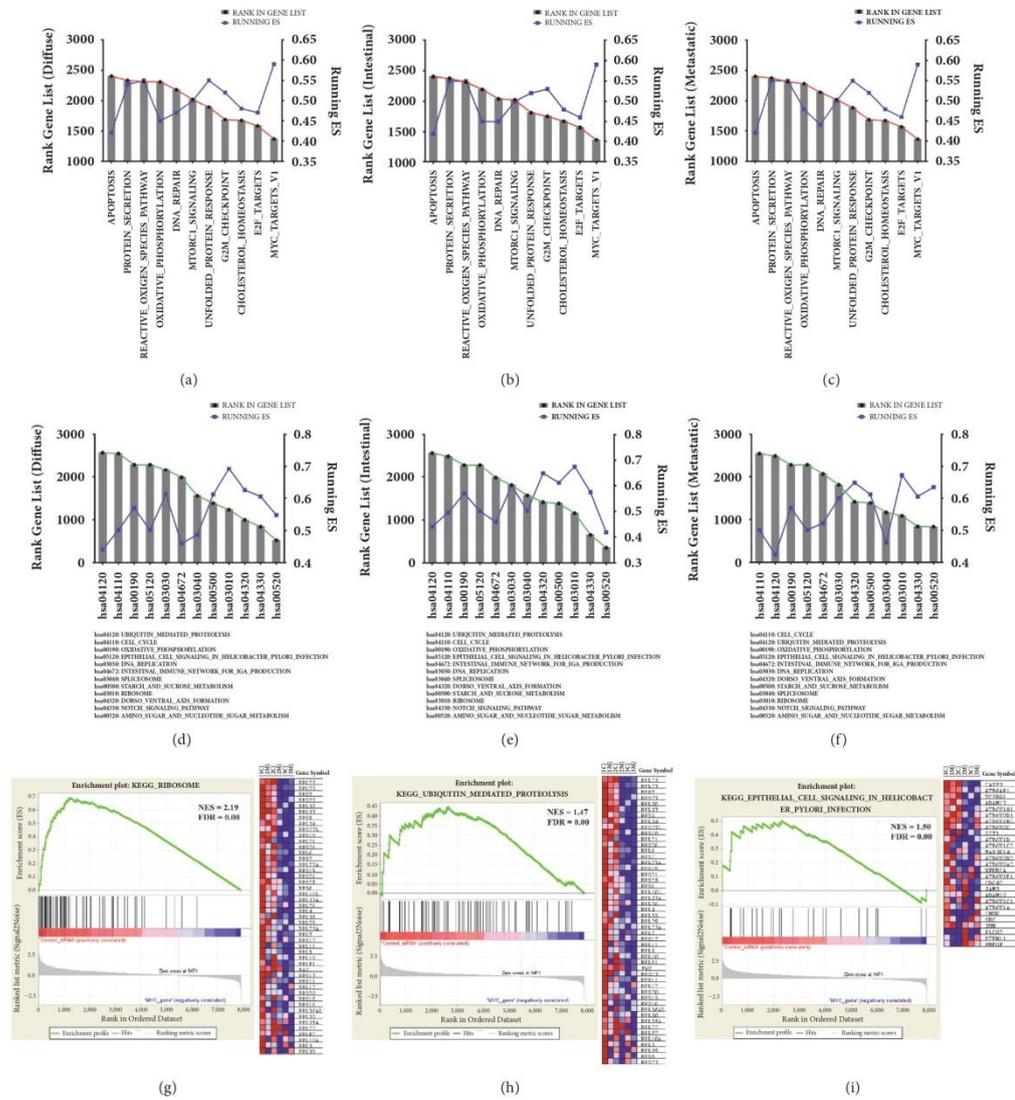


FIGURE 4: Identification of individual genes in the enriched DEG datasets with the higher ES. (a) DEGs enrichment for ACP02. (b) DEGs enrichment for ACP03. (c) DEGs enrichment for AGP01. The most enriched hallmark for the three cell lines as MYC_Target_V1 with an ES = 0.59. (d) KEGG pathway identification for ACP02. (e) KEGG pathway identification for ACP03. (f) KEGG pathway identification for AGP01. Enplot and heatmap for the pathways with the highest ES scores among the three GC cell lines can be seen in (g) for ACP02, in (h) for ACP03, and in (i) for AGP01.

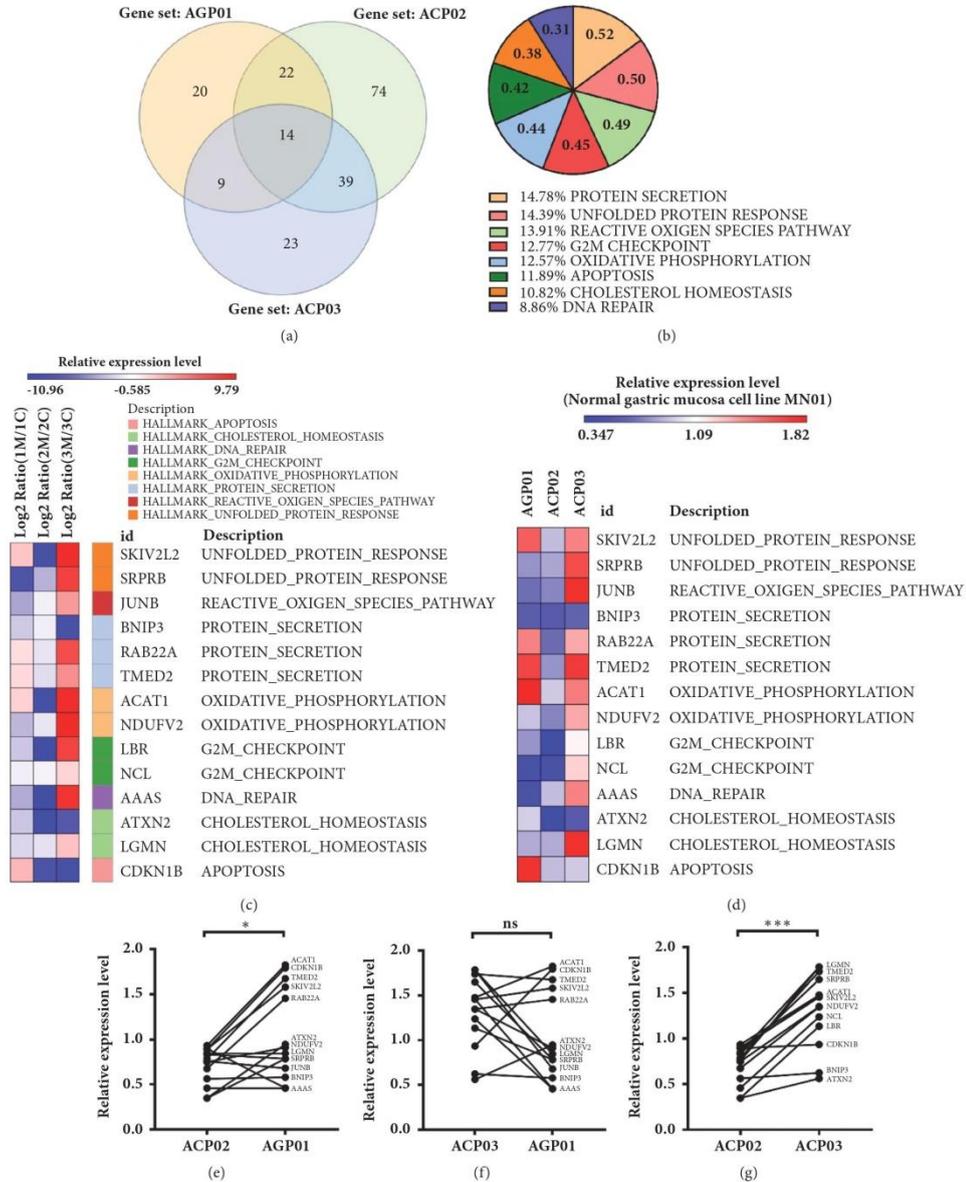


FIGURE 5: Common DEGs enriched in three GC cell lines after MYC-silencing. (a) Using Venn diagram, we were able to identify 14 common DEGs among AGP01, ACP02, and ACP03, but each one of them also has a unique set of DEGs. (b) Identification of the 14 common DEGs in 8 enriched hallmarks of collection H, represented by the % of its correspondent ES, ES scores are shown inside the pie chart for each hallmark. (c) Heatmap for the expression of the 14 common DEGs shared for three GC cell lines after MYC-silencing. Blue represents downregulation while red means upregulation. Notice that, even for the 14 common genes, each cell line has a different expression pattern. Relative read expression was normalized using log₂ Fold-change between MYC-silenced/control siRNA. (d) Heatmap for the relative gene expression obtained by RT-qPCR for the 14 identified DEGs for the AGP01, ACP02, and ACP03 cell lines. (e) The overall expression for the 14 DEGs is downregulated in the ACP02 cell line when compared to AGP01 cells. (f) The gene expression levels were increased in ACP03 cell line, but no statistical difference was found. (g) Relative gene expression was increased for ACP03 when compared to ACP02. Wilcoxon matched-pairs signed rank test was used to compare the relative gene expression levels (*P < 0.05, ns: not significant, and ***P < 0.0001).

DEGs as well as the gene expression hallmarks they are involved, ranks and ES. We show a heatmap (Figure 5(c)) of the 14 shared DEGs in clusters grouped using the GENE E tool (<https://software.broadinstitute.org/GENE-E/>), where downregulation is expressed in blue and upregulation in red. Most genes for ACP02 (diffuse-type GC) presented downregulation, while for ACP03 (intestinal-type GC) the same genes presented themselves as upregulated. The metastatic cell line AGP01, even though its original tumor was intestinal-type, presented mixed expression patterns with predominant downregulation.

3.3. The 14 MYC-Regulated DEGs Show Distinct Expression Profiles for Each GC Histological Subtype. Our *in silico* analysis identified a core set of 14 DEGs (Figure 5(a)) who are MYC-regulated, but whose expression profile is distinct for each GC cell line since they represent different histological subtypes. We validated the gene expression profiles presented by the AGP01, ACP02, and ACP03 after MYC-silencing by RT-qPCR in the cDNA obtained originally for each cell line and compared them with their expression in the MN01 cell line. Our results point out that the 14 identified DEGs are under MYC transcriptional regulation (Figure 5(d)). The ACP02 cell line presented mostly downregulation for the expression of the 14 DEGs (Figure 5(c)); on the other hand, ACP03, the same 14 genes, were upregulated; AGP01 results were mixed, with some genes showing downregulation and others showing upregulation. Our gene expression results (Figure 5(d)) confirm our *in silico* analysis (Figure 5(c)).

Each cell line used in this study represents a GC histological subtype: AGP01 was obtained from the ascitic fluid of intestinal-type GC, representing a metastatic disease, while ACP02 was developed from a diffuse-type stomach cancer patient and ACP03 origin was an intestinal-type gastric tumor.

We compared the mRNA expression measured by RT-qPCR for our identified DEGs to assess whether the gene expression of those 14 genes was enough to statistically distinguish each cell line. When comparing ACP02 versus AGP01 (Figure 5(e)), we noticed a significant gene expression downregulation for ACP02; confronting ACP03 versus AGP01 indicated an increase in mRNA relative quantification (Figure 5(f)) for ACP03; however, those results were not significant; the comparison between ACP02 versus ACP03 confirmed that the 14 MYC-regulated DEGs are significantly more expressed in the ACP03, the intestinal-type cell line, than in ACP02 (Figure 5(g)).

Taken together, our results indicate that, even though MYC-related carcinogenesis alters the same 14 genes in GC cell lines representing the most common histological subtypes, how MYC causes GC carcinogenesis is different for each disease presentation and it is possible to distinct them by using expression signatures.

4. Discussion

A key goal of cancer studies is to systematically characterize the cellular and molecular mechanisms involved in the disease and its distinct stages, to identify both potential

biomarkers and new probable drug targets [44]. The molecular profile of gastric cancer is heterogeneous, partly due to different classification systems [45] and, in order to clarify the true molecular origins of GC, both the Cancer Genome Atlas [22] and the Asian Cancer Research Group [46] published the molecular subtypes of gastric cancer, with remarkable overlap between the two models. Therefore, several genes have been implied as biomarkers for GC subtypes, such as *RHOA*, *EGFR*, *PDL*, *CDH1*, *TP53*, and *JAK2*. However, those studies use samples from populations in which the disease incidence is highest, and few studies have examined populations in which the incidence of this disease is lower, such as Brazil [47]. There is evidence that GC incidence varies between countries greatly because the genetic heterogeneity exhibited by human populations [48], and it has already been showed that there is a unique gene expression signature for Brazilian cases of intestinal-type GC [47]. Our study helps to highlight the molecular profiles of Brazilian GC cell lines, which can help greatly our understanding about the molecular basis of GC in South America.

Most NGS studies investigate GC by comparing tumor versus nontumoral tissue, analyzing global gene expression patterns, copy number variation, and other molecular characteristics, and most of the high-throughput studies carried so far concerning gastric cancer genetics overlook MYC's importance in this process [2, 3, 21–24]. Our results are relevant because MYC overexpression is a key finding in Brazilian GC samples [8]. Therefore, we reduced the expression of this gene using siRNA to identify the MYC-related signature in GC cell lines, comparing nonsilenced with silenced samples. We identified a total of 5,471 DEGs, and 11 significant gene sets, including classic MYC targets represented by MYC target V1.

We hereby present the computational analysis of gene sets identified after MYC-silencing in Brazilian GC cell lines [13, 27], who carry genetic alterations commonly found in Brazilian GC patients [7, 13, 16]. This oncogene promotes cell growth acting as a transcription factor regulating cell cycle, metabolism, and cell survival [49]. We found DEGs upregulation only for the intestinal-type cell line (ACP03), while the diffuse-type (ACP02) and the metastatic GC cells (AGP01) presented overall gene expression downregulation; when looking at individual genes between ACP02 and AGP01, it is still possible to distinguish between them by MYC-related gene expression. It is important to notice that MYC has a dual-role in the carcinogenic process, selectively activating and inactivating different gene sets [50–52]. Taken together, we present evidences supporting the fact that MYC deregulation has an important role in gastric carcinogenesis [14] and that MYC-related signatures in gastric cancer are different for each histological subtype of this disease, which is clinically relevant [47, 53].

One of the main forms of MYC protein regulation in normal cells is through its targeted degradation by the ubiquitin-proteasome system [54], which was one of the KEGG enriched pathways found in our analysis (Figure 4). This means that, in a MYC-overexpression condition, like GC [7], not only this gene and its protein are more produced, but they are also less destroyed because it diminishes the expression of E3 ubiquitin-ligases, such as Fbw7 and HectH9,

contributing to prolonged MYC protein half-life and amplification of its effects [55]. Ubiquitin-ligases, including the MYC-regulated Fbw7, have recently evolved as promising therapeutic targets for the development of novel anticancer drugs [56].

Other pathways involved in MYC-related gastric carcinogenesis found by our study are known targets, such as ribosome and cell cycle control genes, which are hardly druggable. When taken together, the 12 different pathways we found under MYC control for gastric carcinogenesis represent many biological functions, meaning that MYC overexpression in GC disturbs almost all the regular cellular processes in favor of tumor development [52]. Another interesting gene set includes glucose metabolism, which is an area of growing interest in cancer research [57], and it has been shown that MYC directs the activation of aerobic glycolysis, a hallmark of cancer metabolism known as Warburg effect, and pretty much all genes involved in glycolysis and most of the ones responsible for glutaminolysis [57, 58].

We were also able to pinpoint 14 enriched DEGs in all the three GC cell lines used in this study that might represent the common set of MYC-regulated genes in gastric carcinogenesis (Table 4). This is important because it shows that, even though we have represented distinct GC histological subtypes and disease stages, there is still a core set of genes regulated by MYC involved in the carcinogenic process. We also did not find other reports in the literature concerning those 14 genes and high-throughput analysis of gastric cancer [1, 22, 26, 44–46, 53, 59–64]. Even when we consider the molecular signatures presented by Brazilian intestinal-type GC [47], we could not find any concordance for the 14 genes found by our study, but it is important to consider that Binato et al. [47] did not take into account MYC overexpression in their samples. Therefore, the unique DEGs found in this paper represent new and important findings concerning the process of gastric carcinogenesis regulated by MYC in the Brazilian population.

Even though additional studies are needed to validate our results, we present strong evidence that MYC-regulated genes in GC have different expression patterns when we consider histological and disease stage differences; however, they still share pathways and core genes involved in the carcinogenic process.

Data Availability

The gene expression data used to support the findings of this study have been deposited in the Gene Expression Omnibus (GEO) repository under the access number GSE81265 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE81265>). A detailed description of each gene set can be found within the paper at Table 1.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Authors' Contributions

Jersey Heitor da S. Maués, Helem Ferreira Ribeiro, Caroline de Fátima Aquino Moreira-Nunes, and Giovanni R. Pinto conceived and designed the study. Helem Ferreira Ribeiro, Luana de Oliveira Lopes, Giovanni R. Pinto, Letícia M. Lamarão, and Carla Mariana F. Pessoa conducted the study and participated in laboratory analysis. Raimundo Miranda de Carvalho, Jersey Heitor da S. Maués, Helem Ferreira Ribeiro, Juan A. Rey, and Paulo P. Assumpção conducted the RT-qPCR analysis. Jersey Heitor da S. Maués, Caroline de Fátima Aquino Moreira-Nunes, and Rommel M. Rodríguez Burbano conducted bioinformatic analysis and analyzed the data. Jersey Heitor da S. Maués, Helem Ferreira Ribeiro, and Rommel M. Rodríguez Burbano wrote the paper. All authors read and approved the final manuscript.

Acknowledgments

This work was supported by CNPq, Process no. 402283/2013 and Process no. 471072/2012-5; by FAPESPA, Process no. ICAAF 123/2014; by CAPES, Process no. 17704/12-0.

Supplementary Materials

Table 1: enriched hallmark gene sets for DEGs. Table 2: enriched KEGG gene sets for DEGs. Table 3: the 14 common DEGs for three GC cell lines after MYC siRNA and the gene set hallmarks they were enriched and identified in cancer-associated studies. (*Supplementary Materials*)

References

- [1] B. W. Katona and A. K. Rustgi, "Gastric cancer genomics: advances and future directions," *Cellular and Molecular Gastroenterology and Hepatology*, vol. 3, no. 2, pp. 211–217, 2017.
- [2] M. H. McLean and E. M. El-Omar, "Genetics of gastric cancer," *Nature Reviews Gastroenterology & Hepatology*, vol. 11, no. 11, pp. 664–674, 2014.
- [3] K. Wang, S. T. Yuen, J. Xu et al., "Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer," *Nature Genetics*, vol. 46, no. 6, pp. 573–582, 2014.
- [4] A. D. Panani, "Cytogenetic and molecular aspects of gastric cancer: Clinical implications," *Cancer Letters*, vol. 266, no. 2, pp. 99–115, 2008.
- [5] X. Liu, H. Cai, H. Huang et al., "The prognostic significance of apoptosis-related biological markers in chinese gastric cancer patients," *PLoS ONE*, vol. 6, no. 12, 2011.
- [6] J.-S. Choi, J. Seo, E. J. Jung, E. J. Kim, G. K. Lee, and W. H. Kim, "c-MYC amplification in mucinous gastric carcinoma: A possible genetic alteration leading to deeply invasive tumors," *Anticancer Research*, vol. 32, no. 11, pp. 5031–5038, 2012.
- [7] D. Q. Calcagno, V. M. Freitas, M. F. Leal et al., "MYC, FBXW7 and TP53 copy number variation and expression in Gastric Cancer," *BMC Gastroenterology*, vol. 13, no. 1, 2013.
- [8] M. F. Leal, H. F. Ribeiro, J. A. Rey et al., "YWHAE silencing induces cell proliferation, invasion and migration through the up-regulation of CDC25B and MYC in gastric cancer cells: New insights about YWHAE role in the tumor development and

- metastasis process," *Oncotarget*, vol. 7, no. 51, pp. 85393–85410, 2016.
- [9] A. A. Mello, M. F. Leal, J. A. Rey et al., "Deregulated expression of SRC, LYN and CKB kinases by DNA methylation and its potential role in gastric cancer invasiveness and metastasis," *PLoS ONE*, vol. 10, no. 10, 2015.
- [10] W. P. Tansey, "Mammalian MYC Proteins and Cancer," *New Journal of Science*, vol. 2014, Article ID 757534, 27 pages, 2014.
- [11] L. C. Costa Raiol, E. C. Figueira Silva, D. Mendes da Fonseca et al., "Interrelationship between MYC gene numerical aberrations and protein expression in individuals from northern Brazil with early gastric adenocarcinoma," *Cancer Genetics and Cytogenetics*, vol. 181, no. 1, pp. 31–35, 2008.
- [12] D. Q. Calcagno, M. F. Leal, S. Demachki et al., "MYC in gastric carcinoma and intestinal metaplasia of young adults," *Cancer Genetics and Cytogenetics*, vol. 202, no. 1, pp. 63–66, 2010.
- [13] M. F. Leal, D. Q. Calcagno et al., "MYC, TP53, and chromosome 17 copy-number alterations in multiple gastric cancer cell lines and in their parental primary tumors," *Journal of Biomedicine and Biotechnology*, vol. 2011, Article ID 631268, 8 pages, 2011.
- [14] C. R. T. De Souza et al., "MYC deregulation in gastric cancer and its clinicopathological implications," *PLoS One*, vol. 8, no. 5, 2013.
- [15] J. D. F. F. Borges Da Costa et al., "Experimental gastric carcinogenesis in cebus apella nonhuman primates," *PLoS One*, vol. 6, no. 7, p. 13, 2011.
- [16] F. Wisniewski, D. Q. Calcagno, M. F. Leal et al., "Reference genes for quantitative RT-PCR data in gastric tissues and cell lines," *World Journal of Gastroenterology*, vol. 19, no. 41, pp. 7121–7128, 2013.
- [17] J. R. Whitfield and L. Soucek, "Tumor microenvironment: Becoming sick of Myc," *Cellular and Molecular Life Sciences*, vol. 69, no. 6, pp. 931–934, 2012.
- [18] S. A. Hills and J. F. X. Diffley, "DNA replication and oncogene-induced replicative stress," *Current Biology*, vol. 24, no. 10, pp. R435–R444, 2014.
- [19] L. Zhang, Y. Hou, H. Ashktorab et al., "The impact of C-MYC gene expression on gastric cancer cell," *Molecular and Cellular Biochemistry*, vol. 344, no. 1–2, pp. 125–135, 2010.
- [20] S. Walz and M. Eilers, "Unlocking the mysterious mechanisms of Myc," *Nature Medicine*, vol. 19, no. 1, pp. 26–27, 2013.
- [21] P. Karimi, F. Islami, S. Anandasabapathy, N. D. Freedman, and F. Kamangar, "Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention," *Cancer Epidemiology Biomarkers & Prevention*, vol. 23, no. 5, pp. 700–713, 2014.
- [22] A. J. Bass et al., "Comprehensive molecular characterization of gastric adenocarcinoma," *Nature*, vol. 513, no. 7517, pp. 202–209, 2014.
- [23] C. Figueiredo, M. C. Camargo, M. Leite, E. M. Fuentes-Pananá, C. S. Rabkin, and J. C. Machado, *Pathogenesis of Gastric Cancer: Genetics and Molecular Classification BT - Molecular Pathogenesis and Signal Transduction by Helicobacter pylori*, N. Tegtmeyer and S. Backert, Eds., vol. 400 of *Current Topics in Microbiology and Immunology*, Springer International Publishing, 2017.
- [24] J. Cools-Lartigue, L. Baker, and L. E. Ferri, *Molecular Mechanisms in Gastric Carcinogenesis BT - Gastric Cancer: Principles and Practice*, V. E. Strong, Ed., Springer International Publishing, 2015.
- [25] M. Kato, "Bioinformatics in cancer clinical sequencing—an emerging field of cancer personalized medicine," *Gan To Kagaku Ryoho*, vol. 43, no. 4, pp. 391–397, 2016.
- [26] C. Sun, Q. Yuan, D. Wu, X. Meng, and B. Wang, "Identification of core genes and outcome in gastric cancer using bioinformatics analysis," *Oncotarget*, vol. 8, no. 41, pp. 70271–70280, 2017.
- [27] M. F. Leal, J. L. Martins do Nascimento, C. E. A. da Silva et al., "Establishment and conventional cytogenetic characterization of three gastric cancer cell lines," *Cancer Genetics and Cytogenetics*, vol. 195, no. 1, pp. 85–91, 2009.
- [28] A. S. Khayat, A. C. Guimarães, D. Q. Calcagno et al., "Interrelationship between TP53 gene deletion, protein expression and chromosome 17 aneusomy in gastric adenocarcinoma," *BMC Gastroenterology*, vol. 9, no. 1, 2009.
- [29] H. F. Ribeiro, D. F. A. Alcântara, L. A. Matos et al., "Cytogenetic characterization and evaluation of c-MYC gene amplification in PG100, a new Brazilian gastric cancer cell line," *Brazilian Journal of Medical and Biological Research*, vol. 43, no. 8, pp. 717–721, 2010.
- [30] H. Li and R. Durbin, "Fast and accurate long-read alignment with Burrows-Wheeler transform," *Bioinformatics*, vol. 26, no. 5, pp. 589–595, 2010.
- [31] N. Leng, J. A. Dawson, J. A. Thomson et al., "EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments," *Bioinformatics*, vol. 29, no. 8, pp. 1035–1043, 2013.
- [32] H. Li, "Exploring single-sample snp and indel calling with whole-genome de novo assembly," *Bioinformatics*, vol. 28, no. 14, Article ID bts280, pp. 1838–1844, 2012.
- [33] R. Patro, S. M. Mount, and C. Kingsford, "Sailfish enables alignment-free isoform quantification from RNA-seq reads using lightweight algorithms," *Nature Biotechnology*, vol. 32, no. 5, pp. 462–464, 2014.
- [34] C. Trapnell, B. A. Williams, G. Pertea et al., "Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation," *Nature Biotechnology*, vol. 28, no. 5, pp. 511–515, 2010.
- [35] M. Garber, M. G. Grabherr, M. Guttman, and C. Trapnell, "Computational methods for transcriptome annotation and quantification using RNA-seq," *Nature Methods*, vol. 8, no. 6, pp. 469–477, 2011.
- [36] S. Audic and J.-M. Claverie, "The significance of digital gene expression profiles," *Genome Research*, vol. 7, no. 10, pp. 986–995, 1997.
- [37] H. Abdi, "The Bonferroni and Šidák Corrections for Multiple Comparisons," pp. 1–9, 2007.
- [38] Y. Benjamini and D. Yekutieli, "The control of the false discovery rate in multiple testing under dependency," *The Annals of Statistics*, vol. 29, no. 4, pp. 1165–1188, 2001.
- [39] A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, and B. L. Ebert, "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide," 2005.
- [40] A. Liberzon, C. Birger, H. Thorvaldsdóttir, M. Ghandi, J. P. Mesirov, and P. Tamayo, "The molecular signatures database hallmark gene set collection," *Cell Systems*, vol. 1, no. 6, pp. 417–425, 2015.
- [41] A. Liberzon, A. Subramanian, R. Pinchback, H. Thorvaldsdóttir, P. Tamayo, and J. P. Mesirov, "Molecular signatures database (MSigDB) 3.0," *Bioinformatics*, vol. 27, no. 12, pp. 1739–1740, 2011.
- [42] M. Kanehisa, S. Goto, Y. Sato, M. Furumichi, and M. Tanabe, "KEGG for integration and interpretation of large-scale molecular data sets," *Nucleic Acids Research*, vol. 40, no. 1, pp. D109–D114, 2012.
- [43] H. Heberle, V. G. Meirelles, F. R. da Silva, G. P. Telles, and R. Minghim, "InteractiVenn: a web-based tool for the analysis of

- sets through Venn diagrams," *BMC Bioinformatics*, vol. 16, no. 1, article 169, 2015.
- [44] J. Y. Cho, "Molecular diagnosis for personalized target therapy in gastric cancer," *Gastric Cancer*, vol. 13, no. 3, pp. 129–135, 2013.
- [45] M. Rugge, M. Fassan, and D. Y. Graham, "Epidemiology of gastric cancer," *Gastric Cancer: Principles and Practice*, pp. 23–34, 2015.
- [46] R. Cristescu, J. Lee, M. Nebozhyn et al., "Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes," *Nature Medicine*, vol. 21, no. 5, pp. 449–456, 2015.
- [47] R. Binato, E. C. Santos, M. Boroni, S. Demachki, P. Assumpção, and E. Abdelhay, "A common molecular signature of intestinal-type gastric carcinoma indicates processes related to gastric carcinogenesis," *Oncotarget*, vol. 9, no. 7, pp. 7359–7371, 2018.
- [48] T. E. Buffart, M. Louw, N. C. van Grieken et al., "Gastric cancers of Western European and African patients show different patterns of genomic instability," *BMC Medical Genomics*, vol. 4, no. 1, 2011.
- [49] V. Posternak and M. D. Cole, "Strategically targeting MYC in cancer," *F1000Research*, vol. 5, 2016.
- [50] S. Walz, F. Lorenzin, J. Morton et al., "Activation and repression by oncogenic MYC shape tumour-specific gene expression profiles," *Nature*, vol. 511, no. 7510, pp. 483–487, 2014.
- [51] T. R. Kress, A. Sabò, and B. Amati, "MYC: Connecting selective transcriptional control to global RNA production," *Nature Reviews Cancer*, vol. 15, no. 10, pp. 593–607, 2015.
- [52] T. R. Kress, P. Pellanda, L. Pellegrinet et al., "Identification of MYC-dependent transcriptional programs in oncogene-addicted liver tumors," *Cancer Research*, vol. 76, no. 12, pp. 3463–3472, 2016.
- [53] S. Zang, R. Guo, R. Xing et al., "Identification of differentially-expressed genes in intestinal gastric cancer by microarray analysis," *Genomics, Proteomics & Bioinformatics*, vol. 12, no. 6, pp. 276–283, 2014.
- [54] A. S. Farrell and R. C. Sears, "MYC degradation," *Cold Spring Harbor Perspectives in Medicine*, vol. 4, no. 3, Article ID a014365, 2014.
- [55] B. King, T. Trimarchi, L. Reavie et al., "The ubiquitin ligase FBXW7 modulates leukemia-initiating cell activity by regulating MYC stability," *Cell*, vol. 153, no. 7, pp. 1552–1566, 2013.
- [56] J. Liu, S. Shaik, X. Dai et al., "Targeting the ubiquitin pathway for cancer treatment," *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, vol. 1855, no. 1, pp. 50–60, 2015.
- [57] C. V. Dang, "MYC, metabolism, cell growth, and tumorigenesis," *Cold Spring Harbor Perspectives in Medicine*, vol. 3, no. 8, 2013.
- [58] Z. E. Stine, Z. E. Walton, B. J. Altman, A. L. Hsieh, and C. V. Dang, "MYC, metabolism, and cancer," *Cancer Discovery*, vol. 5, no. 10, pp. 1024–1039, 2015.
- [59] S. S. Wong, K. Kim, J. C. Ting et al., "Genomic landscape and genetic heterogeneity in gastric adenocarcinoma revealed by whole-genome sequencing," *Nature Communications*, vol. 5, no. 1, 2014.
- [60] S. K. Garattini, D. Basile, M. Cattaneo et al., "Molecular classifications of gastric cancers: Novel insights and possible future applications," *World Journal of Gastrointestinal Oncology*, vol. 9, no. 5, pp. 194–208, 2017.
- [61] E. Bria, S. Pilotto, M. Simbolo et al., "Comprehensive molecular portrait using next generation sequencing of resected intestinal-type gastric cancer patients dichotomized according to prognosis," *Scientific Reports*, vol. 6, no. 1, 2016.
- [62] Y. Lin, Z. Wu, W. Guo, and J. Li, "Gene mutations in gastric cancer: a review of recent next-generation sequencing studies," *Tumor Biology*, vol. 36, no. 10, pp. 7385–7394, 2015.
- [63] J. Cools-Lartigue, L. Baker, and L. E. Ferri, "Molecular mechanisms in gastric carcinogenesis," *Gastric Cancer: Principles and Practice*, pp. 35–56, 2015.
- [64] M. Kanda and Y. Kodera, "Recent advances in the molecular diagnostics of gastric cancer," *World Journal of Gastroenterology*, vol. 21, no. 34, pp. 9838–9852, 2015.

APÊNDICE B – CONJUNTO DE GENES MARCADORES ENRIQUECIDOS POR DEGs

ACP02: Difuso				
Conjunto de Genes	Símbolo Gene	Classificação na Lista de Gene	Classificação de Pontuação Métrica	Corrida (ES)
ALVOS <i>MYC</i> V1	<i>SNRPD2</i>	1.368	6.2E-01	0.59
VIA DE ESPÉCIES REATIVAS DE OXIGÊNIO	<i>GPX3</i>	2.316	4.2E-01	0.55
RESPOSTA A PROTEÍNAS MAL ENOVELADAS	<i>NFYB</i>	1.891	4.9E-01	0.55
SECREÇÃO PROTEICA	<i>RER1</i>	2.334	4.2E-01	0.54
G2M_CHECKPOINT	<i>RBM14</i>	1.686	5.4E-01	0.52
SINALIZAÇÃO MTORC1	<i>YKT6</i>	2.021	4.7E-01	0.50
HOMEOSTASE DE COLESTEROL	<i>EBP</i>	1.671	5.4E-01	0.48
FOSFORILAÇÃO OXIDATIVA	<i>ATP5F1</i>	2.181	4.4E-01	0.47
ALVOS E2F	<i>DEPDC1</i>	1.582	5.6E-01	0.47
REPARO DE DNA	<i>NME3</i>	2.311	4.2E-01	0.45
APOPTOSE	<i>CDKN1B</i>	2.405	4.0E-01	0.42
ACP03: Intestinal				
Conjunto de Genes	Símbolo Gene	Classificação na Lista de Gene	Classificação de Pontuação Métrica	Corrida (ES)
ALVOS <i>MYC</i> V1	<i>TYMS</i>	1.365	6.2E-01	0.59
SECREÇÃO PROTEICA	<i>KIF1B</i>	2.376	4.1E-01	0.55
VIA DE ESPÉCIES REATIVAS DE OXIGÊNIO	<i>GPX3</i>	2.316	4.2E-01	0.55
RESPOSTA A PROTEÍNAS MAL ENOVELADAS	<i>EDEMI</i>	1.752	5.3E-01	0.53
G2M_CHECKPOINT	<i>DBF4</i>	1.813	5.1E-01	0.52
SINALIZAÇÃO MTORC1	<i>YKT6</i>	2.021	4.7E-01	0.50
HOMEOSTASE DE COLESTEROL	<i>EBP</i>	1.671	5.4E-01	0.48
ALVOS E2F	<i>TIMELESS</i>	1.568	5.7E-01	0.46
REPARO DE DNA	<i>RNMT</i>	2.194	4.4E-01	0.45
FOSFORILAÇÃO OXIDATIVA	<i>NDUFS8</i>	2.038	4.7E-01	0.45
APOPTOSE	<i>CDKN1B</i>	2.405	4.0E-01	0.42
AGP01: Líquido ascítico do intestino (Metastático)				
Conjunto de Genes	Símbolo Gene	Classificação na Lista de Gene	Classificação de Pontuação Métrica	Corrida (ES)
ALVOS <i>MYC</i> V1	<i>TYMS</i>	1.365	6.2E-01	0.59
SECREÇÃO PROTEICA	<i>KIF1B</i>	2.376	4.1E-01	0.55
VIA DE ESPÉCIES REATIVAS DE OXIGÊNIO	<i>GPX3</i>	2.316	4.2E-01	0.55
RESPOSTA A PROTEÍNAS MAL ENOVELADAS	<i>NFYB</i>	1.891	4.9E-01	0.55
G2M CHECKPOINT	<i>RBM14</i>	1.686	5.4E-01	0.52
SINALIZAÇÃO MTORC1	<i>YKT6</i>	2.021	4.7E-01	0.50
FOSFORILAÇÃO OXIDATIVA	<i>ACADSB</i>	2.284	4.3E-01	0.48
HOMEOSTASE DE COLESTEROL	<i>EBP</i>	1.671	5.4E-01	0.48
ALVOS E2F	<i>TIMELESS</i>	1.568	5.7E-01	0.46
REPARO DE DNA	<i>ERCC3</i>	2.142	4.5E-01	0.44
APOPTOSE	<i>CDKN1B</i>	2.405	4.0E-01	0.42

ES: escore de enriquecimento.

APÊNDICE C – CONJUNTO DE GENES KEGG ENRIQUECIDOS POR DEGs

ACP02: Diffuse				
Conjuntos de genes KEGG	Símbolo Gene	Classificação na Lista de Gene	Classificação de Pontuação Métrica	Corrida (ES)
RIBOSSOMO	<i>RPS15</i>	1.233	7.1E-01	0.69
FORMAÇÃO DO EIXO VENTRAL DE DORSO	<i>SOS2</i>	996	6.0E-01	0.62
REPLICAÇÃO DE DNA	<i>RNASEH2B</i>	2.165	4.5E-01	0.61
METABOLISMO DE AMIDO E SACAROSE	<i>PYGL</i>	1.386	6.1E-01	0.61
VIA DE SINALIZAÇÃO DE NOTCH	<i>DVL3</i>	837	8.1E-01	0.61
FOSFORILAÇÃO OXIDATIVA	<i>ATP6V0A2</i>	2.289	4.3E-01	0.57
METABOLISMO NUCLEOTÍDEO DE AMINO E AÇÚCAR	<i>AMDHD2</i>	521	1.0E+16	0.55
SINALIZAÇÃO CELULAR EPITELIAL EM INFECÇÃO POR <i>H. pylori</i>	<i>ATP6V0A2</i>	2.289	4.3E-01	0.50
CICLO CELULAR	<i>CDK4</i>	2.549	3.8E-01	0.50
SPLICEOSOMA	<i>PRPF18</i>	1.550	5.7E-01	0.49
REDE IMUNOLÓGICA INTESTINAL PARA PRODUÇÃO IGA	<i>MAP3K14</i>	1.984	4.8E-01	0.46
PROTEÍLISE MEDIADA UBIQUITINA	<i>UBE2J2</i>	2.568	3.7E-01	0.44
ACP03: Intestinal				
Conjuntos de genes KEGG	Símbolo Gene	Classificação na Lista de Gene	Classificação de Pontuação Métrica	Corrida (ES)
RIBOSSOMO	<i>RPL12</i>	1.159	6.8E-01	0.68
FORMAÇÃO DO EIXO VENTRAL DE DORSO	<i>SOS1</i>	1.415	6.0E-01	0.65
METABOLISMO DE AMIDO E SACAROSE	<i>PYGL</i>	1.386	6.1E-01	0.61
REPLICAÇÃO DE DNA	<i>RNASEH2C</i>	1.809	5.1E-01	0.60
VIA DE SINALIZAÇÃO DE NOTCH	<i>NOTCH2</i>	654	9.0E-01	0.57
FOSFORILAÇÃO OXIDATIVA	<i>ATP6V0A2</i>	2.289	4.3E-01	0.57
SPLICEOSOMA	<i>DHX16</i>	1.569	5.7E-01	0.50
SINALIZAÇÃO CELULAR EPITELIAL EM INFECÇÃO POR <i>H. pylori</i>	<i>ATP6V0A2</i>	2.289	4.3E-01	0.50
CICLO CELULAR	<i>ANAPC1</i>	2.493	3.9E-01	0.50
REDE IMUNOLÓGICA INTESTINAL PARA PRODUÇÃO IGA	<i>MAP3K14</i>	1.984	4.8E-01	0.46
PROTEÍLISE MEDIADA UBIQUITINA	<i>UBE2J2</i>	2.568	3.7E-01	0.44
METABOLISMO NUCLEOTÍDEO DE AMINO E AÇÚCAR	<i>GPI</i>	354	1.2E+16	0.42
AGP01: Líquido ascítico do intestino (Metastático)				
Conjuntos de genes KEGG	Símbolo Gene	Classificação na Lista de Gene	Classificação de Pontuação Métrica	Corrida (ES)
RIBOSSOMO	<i>RPS11</i>	1.086	7.1E-01	0.67
FORMAÇÃO DO EIXO VENTRAL DE DORSO	<i>SOS1</i>	1.415	7.4E-01	0.65
METABOLISMO NUCLEOTÍDEO DE AMINO E AÇÚCAR	<i>GNPDA1</i>	831	8.1E-01	0.63
METABOLISMO DE AMIDO E SACAROSE	<i>PYGL</i>	1.386	6.1E-01	0.61
VIA DE SINALIZAÇÃO DE NOTCH	<i>DVL3</i>	837	8.1E-01	0.61
REPLICAÇÃO DE DNA	<i>RNASEH2C</i>	1.809	5.1E-01	0.60
FOSFORILAÇÃO OXIDATIVA	<i>ATP6V0A2</i>	2.289	4.3E-01	0.57
REDE IMUNOLÓGICA INTESTINAL PARA PRODUÇÃO IGA	<i>IL6</i>	2.064	4.7E-01	0.52
SINALIZAÇÃO CELULAR EPITELIAL EM INFECÇÃO POR <i>H. pylori</i>	<i>ATP6V0A2</i>	2.289	4.4E-01	0.50
CICLO CELULAR	<i>CDK4</i>	2.549	3.8E-01	0.50
SPLICEOSOMA	<i>SNRPF</i>	1.167	6.8E-01	0.46
PROTEÍLISE MEDIADA POR UBIQUITINA	<i>ANAPC1</i>	2.493	3.9E-01	0.42

ES: escore de enriquecimento.

APÊNDICE D - Os 14 DEGs comuns para as três linhagens celulares de GC após *MYC*-siRNA e as características dos genes que foram enriquecidos e identificados em estudos associados ao câncer.

Nome da característica:	Símbolo Gene	Classificação na Lista de Gene	Classificação de Pontuação Métrica	Corrida (ES)
SECREÇÃO PROTEICA	<i>BNIP3</i> [1]	1438	6.0E-01	0.522
RESPOSTA A PROTEÍNAS MAL ENOVELADAS	<i>SKIV2L2</i> [2]	1051	7.2E-01	0.508
RESPOSTA A PROTEÍNAS MAL ENOVELADAS	<i>SRPRB</i> [3]	1442	6.0E-01	0.506
VIA DE ESPÉCIES REATIVAS DE OXIGÊNIO	<i>JUNB</i> [4]	1969	4.8E-01	0.491
G2M CHECKPOINT	<i>LBR</i> [5]	817	8.2E-01	0.451
FOSFORILAÇÃO OXIDATIVA	<i>NDUFB2</i> [6]	1542	5.7E-01	0.444
APOPTOSE	<i>CDKN1B</i> [7]	2405	4.0E-01	0.420
SECREÇÃO PROTEICA	<i>RAB22A</i> [8]	289	1.2E+15	0.408
HOMEOSTASE DE COLESTEROL	<i>ATXN2</i> [9]	945	7.6E-01	0.382
HOMEOSTASE DE COLESTEROL	<i>LGMN</i> [10]	445	1.1E+16	0.367
FOSFORILAÇÃO OXIDATIVA	<i>ACATI</i> [11]	840	8.1E-01	0.367
SECREÇÃO PROTEICA	<i>TMED2</i> [12]	254	1.3E+16	0.326
REPARO DE DNA	<i>AAAS</i> [13]	1109	7.0E-01	0.313
G2M CHECKPOINT	<i>NCL</i> [14]	328	1.2E+13	0.289

ES: escore de enriquecimento.